

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2006, 18:19:09 ; Search time 215.124 Seconds  
(without alignments)  
4078.762 Million cell updates/sec

Title: US-10-633-742-2  
Perfect score: 10483  
Sequence: 1 MLHGAGLAWITLSLQGT.....FPIYENVPEYHRDPVYSRH 1997

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21.\*  
1: geneseqp1980s.\*  
2: geneseqp1990s.\*  
3: geneseqp2000s.\*  
4: geneseqp2001s.\*  
5: geneseqp2002s.\*  
6: geneseqp2003as.\*  
7: geneseqp2003bs.\*  
8: geneseqp2004s.\*  
9: geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10483	100.0	1997	3	AAB19774 Human pro
2	10483	100.0	1997	8	ADO04580 Human HPT
3	10460	99.8	1997	5	AAE20278 Human lun
4	10460	99.8	1997	7	ABR57179 Human PTP
5	10460	99.8	1997	7	ADJ70329 Human hea
6	10457	99.8	1997	4	AAM78821 Human pro
7	10440.5	99.6	2002	4	AAW79805 Human pro
8	9117.5	87.0	1998	7	AAO24268 Murine va
9	7614	72.6	1450	7	ABR57182 Human PTP
10	6712	64.0	1407	8	ADR10246 Human pro
11	2937	28.0	579	3	AAB19773 Mouse vas
12	1854.5	17.7	609	8	ADQ66828 Novel hum
13	1830	17.5	336	8	ADO04582 Human HPT
14	1691	16.1	319	8	ADO04585 Human HPT
15	1690	16.1	312	8	ADO04584 Human HPT
16	1556.5	14.8	1647	4	ABR58428 Drosophil
17	1485.5	14.2	1767	4	ABB62726 Drosophil
18	1433.5	13.7	1447	7	ABR57181 Drosophil
19	1374	13.1	310	4	AAG78275 Human PTP
20	1367	13.0	254	4	AAR59377 Human pro
21	1247	11.9	1337	2	AAR85203 huDBP-1.
22	1247	11.9	1337	9	ADY37133 Protein t
23	1243.5	11.9	1238	7	ADL16198 Mouse pro
24	1242	11.8	1337	7	ADL16193 Human pro

25	1242	11.8	1337	8	ADP74603 Amino aci
26	1242	11.8	1337	9	ADY37011 Protein t
27	1221.5	11.7	1216	7	ADL16200 Rat prote
28	1206	11.5	1337	7	ADL16195 Human pro
29	1206	11.5	1337	9	ADY37013 Protein t
30	1204	11.5	1705	5	ABR52348 Protein r
31	1198	11.4	1711	2	AAW70506 Osteotest
32	1198	11.4	1711	5	ABR52349 Protein r
33	1188	11.3	1711	2	AAW70507 Mutant os
34	1121	10.7	2299	6	AAE37320 Human rec
35	1119.5	10.7	1959	6	AAE37322 Human rec
36	1114	10.6	2299	8	ADH41615 Novel hum
37	1107	10.6	2291	6	ABP60057 Human pho
38	1106	10.6	2281	5	AAO18736 Human NOV
39	1106	10.6	2281	8	ADH41619 Novel hum
40	1106	10.6	2281	8	ADH41613 Novel hum
41	1100.5	10.5	2300	5	AAO18738 Human NOV
42	1100.5	10.5	2300	8	ADH41625 Novel hum
43	1099.5	10.5	2299	6	AAG79724 Human KPP
44	1097	10.5	2301	6	ABP60058 Human pho
45	1053	10.0	337	7	AAO24271 Murine va

## ALIGNMENTS

RESULT 1  
AAB19774  
ID AAB19774 standard; protein; 1997 AA.  
XX  
AC AAB19774;  
XX  
DT 19-FEB-2001 (first entry)  
XX  
DE Human protein tyrosine phosphatase HPTP-beta.  
XX  
KW Protein tyrosine phosphatase; HPTP-beta; human; Tie-2;  
KW receptor-type tyrosine kinase; antiangiogenic; antitumour;  
KW antimetastatic; tumour; metastasis; angiogenesis; therapy.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain 1622..1641  
FT FT /note= "transmembrane domain"  
FT Domain 1722..1967  
FT FT /note= "catalytic domain"  
XX  
PN EPI046715-A1.

XX 25-OCT-2000.  
XX 23-APR-1999; 99EP-00108074.  
XX 23-APR-1999; 99EP-00108074.  
XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
XX Fachinger G, Rieau B, Deutsch U;  
XX WPI; 2000-648932/63  
XX N-PSDB; AAA88866.  
XX Monitoring of modulating Tie-2 tyrosine kinase activity, useful e.g. for  
XX regulating tumor growth, using vascular-endothelial protein tyrosine  
XX phosphatase.

XX Disclosure; Page 21-27; 60pp; English.

XX The present sequence is that of human protein tyrosine phosphatase HPTP-  
XX beta, a member of subclass III receptor type PTPs, bearing fibronectin  
XX type III-like repeats in the extracellular domain and a single catalytic  
XX domain in the cytoplasmic tail. HPTP-beta is a vascular-endothelial

CC	protein tyrosine phosphatase (VE-PTP) that specifically interacts with									
CC	receptor-type tyrosine kinase Tie-2, modulating its tyrosine									
CC	phosphorylation. Tie-2 is involved in angiogenic processes, the									
CC	formation of blood vessels during embryonal development, wound healing									
CC	and in pathological processes such as tumour development. VE-PTPs such as									
CC	HPTP-beta or its catalytic domain, nucleic acids and ligands can be used									
CC	to monitor, stimulate or repress Tie-2 activity for the purpose of									
CC	monitoring or modulating angiogenesis, inducing or inhibiting vascular									
CC	growth or remodelling and blood vessel maturation, and inhibiting tumour									
CC	growth or metastasis									
XX										
SQ	Sequence 1997 AA;									
	Query Match	100.0%;	Score 10483;	DB 3;	Length 1997;					
	Best Local Similarity	100.0%;	Pred. No. 0;							
	Matches 1997;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;					
Qy	1	MLSHGAGLAWITLSLQTLGLAEPERCNFTLAESKASHSVSIQWRILGSPCNFSLYYSS	60							
Db	1	MLSHGAGLAWITLSLQTLGLAEPERCNFTLAESKASHSVSIQWRILGSPCNFSLYYSS	60							
Qy	61	DTLGAALCPTFRIDNTTYGCNLQDLQAGTIYFKIIISLDEERTVVLQTDPLPPARFGVSK	120							
Db	61	DTLGAALCPTFRIDNTTYGCNLQDLQAGTIYFKIIISLDEERTVVLQTDPLPPARFGVSK	120							
Qy	121	EKTSTGLHWHTPSSGKVTSYEVQLFDENNKIQGVQIOBSTSWNEYTFPNLTAGSKYN	180							
Db	121	EKTSTGLHWHTPSSGKVTSYEVQLFDENNKIQGVQIOBSTSWNEYTFPNLTAGSKYN	180							
Qy	181	IAITAVSGGKRSFSVYTINGSTVSPVKDIGISTKANSLISWSHGSGNVSERYRLMLMDKG	240							
Db	181	IAITAVSGGKRSFSVYTINGSTVSPVKDIGISTKANSLISWSHGSGNVSERYRLMLMDKG	240							
Qy	241	ILVHGGVVDVXHATSYAFHGLSPGYLYNLVTWTEAAGLQNYRWKLVRTAPMEVSNLKVTVND	300							
Db	241	ILVHGGVVDVXHATSYAFHGLSPGYLYNLVTWTEAAGLQNYRWKLVRTAPMEVSNLKVTVND	300							
Qy	301	GSUTSLKVKQWRPPGNVDVSNITLSHGKGTIKESRVLAPWITETHFKELVPGRLYQVTVSC	360							
Db	301	GSUTSLKVKQWRPPGNVDVSNITLSHGKGTIKESRVLAPWITETHFKELVPGRLYQVTVSC	360							
Qy	361	VSGELSAQKMAVGRTFPDPKVANLEANNNGRMRSILVVSWSPPAGDWEQYRILLFNDSVLL	420							
Db	361	VSGELSAQKMAVGRTFPDPKVANLEANNNGRMRSILVVSWSPPAGDWEQYRILLFNDSVLL	420							
Qy	421	NIITVGKEETQYVMDDTGLVPGRQYEVEVIVESGNLKNRSERCCQRTVPLAVLQLRVKHANE	480							
Db	421	NIITVGKEETQYVMDDTGLVPGRQYEVEVIVESGNLKNRSERCCQRTVPLAVLQLRVKHANE	480							
Qy	481	TSLSIMMQTPVAWEKIIISLADBDLLIIHKSLSKDAKETFTDLPVGRKYMATVTSISG	540							
Db	481	TSLSIMMQTPVAWEKIIISLADBDLLIIHKSLSKDAKETFTDLPVGRKYMATVTSISG	540							
Qy	541	DLKNSSSVKGRTPVAQVTDLHVANQGMVTSLSFTNWTQAGDVEPYQVLLIHENVVIKNES	600							
Db	541	DLKNSSSVKGRTPVAQVTDLHVANQGMVTSLSFTNWTQAGDVEPYQVLLIHENVVIKNES	600							
Qy	601	ISSETSRYSFHSLKSGSLYISVVVTVTSGGGISRRQVVVEGRTPVSSVSGVTVNNSGRNDYL	660							
Db	601	ISSETSRYSFHSLKSGSLYISVVVTVTSGGGISRRQVVVEGRTPVSSVSGVTVNNSGRNDYL	660							
Qy	661	SVSWLVA PGDVNDVYEVTLSHDGKVVQSLVITAKSVRECSFSSLTPGRLYTVTITRSKYE	720							
Db	661	SVSWLVA PGDVNDVYEVTLSHDGKVVQSLVITAKSVRECSFSSLTPGRLYTVTITRSKYE	720							
Qy	721	NHSFQSORTVPDKVQGVSVSNSARSDYLRSVSWHATGDFDHYEVTIKKNKNFIQTKSIPK	780							
Db	721	NHSFQSORTVPDKVQGVSVSNSARSDYLRSVSWHATGDFDHYEVTIKKNKNFIQTKSIPK	780							
Qy	781	SENECVFQVLVPGRLYISVTVTTKSGQYEANEGQNGRTIPBPVKDLTLRNSTEDLHWTS	840							
Db	781	SENECVFQVLVPGRLYISVTVTTKSGQYEANEGQNGRTIPBPVKDLTLRNSTEDLHWTS	840							

```

Db 1921 LQQLDSKSDVYIGAVHDLRLHVRVWVQTEQYVYLHQCVRDLVLRKRLSEQNPLPPI 1980
Qy 1981 YENVNPEYHRDPVYSRH 1997
Db 1981 YENVNPEYHRDPVYSRH 1997

RESULT 2
ADO04580
ID ADO04580 standard; protein; 1997 AA.
XX
AC ADO04580;
DT 15-JUL-2004 (first entry)
XX
DE Human HPTPbeta protein.
XX
KW Protein co-ordinate data; HPTPbeta; HPTP-beta; PTPRB; PTPbeta; PTPB;
KW R-PTP-beta; angiogenesis mediated disorder; diabetic retinopathy;
KW sickle cell anaemia; Paget's disease; mycobacterial infection;
KW systemic lupus erythematosus; myopia; Crohn's disease; psoriasis;
KW rheumatoid arthritis; tumour; acquired immune deficiency syndrome; AIDS;
KW drug designing; therapy; human.
XX
OS Homo sapiens.
XX
UN US2004077065-A1.
XX
PR 22-APR-2004.
XX
PR 04-AUG-2003; 2003US-00634027.
XX
PR 25-SEP-2002; 2002US-0413547P.
XX
PA (PROC ) PROCTER & GAMBLE CO.
XX
PI Evdokimov AG, Pokross ME;
XX
DR WPI; 2004-374235/35.
XX
DR N-PSDB; ADO04579.
XX
DR GENBANK; X54131.
XX
PT Identification of compound useful for treatment of angiogenesis mediated
PT disorder, by using three-dimensional structure of HPTPbeta catalytic
PT domain, and employing structure to design, or select compound that binds
PT HPTPbeta in silico.
XX
PS Disclosure; SEQ ID NO 2; 335pp; English.
XX
CC The invention relates to the three dimensional coordinates of HPTPbeta
CC (also known as HPTP-beta, PTPRB, PTPbeta, PTPB or R-PTP-beta) protein. It
CC also relates to a method for the identification of a compound useful for
CC the treatment of an angiogenesis mediated disorder. The compounds
CC identified by this method are useful to treat diseases like diabetic
CC retinopathy, sickle cell anaemia, Paget's disease, mycobacterial
CC infections, systemic lupus erythematosus, myopia, Crohn's disease,
CC psoriasis, rheumatoid arthritis, solid or blood borne tumours and
CC acquired immune deficiency syndrome (AIDS). The invention is useful for
CC the treatment of an angiogenesis mediated disorder or disease. It is also
CC useful in drug design techniques. The present sequence is human HPTPbeta
CC protein.
XX
SQ Sequence 1997 AA;

Query Match 100.0%; Score 10483; DB 8; Length 1997;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1997; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSLHGAGLAWITLSLQTLGAEPRCNCFTLAESKASSHSVSICWRIILGSPCNFSLIYSS 60
Db 1 MSLHGAGLAWITLSLQTLGAEPRCNCFTLAESKASSHSVSICWRIILGSPCNFSLIYSS 60

```

```

Qy 61 DTLAGAALCPTFRIDNTTYGNCNLOAQGTIYNFKIISLDEERTVVLQDPLPPARFGVSK 120
Db 61 DTLAGAALCPTFRIDNTTYGNCNLOAQGTIYNFKIISLDEERTVVLQDPLPPARFGVSK 120
Qy 121 EKTITSTGLHWMTTPSSGKVTSEYVQLPDENNQIKQGVQIQESTSWNEYYTFNLTAGSKYN 180
Db 121 EKTITSTGLHWMTTPSSGKVTSEYVQLPDENNQIKQGVQIQESTSWNEYYTFNLTAGSKYN 180
Qy 181 IAITAVSGGKRSPSVYTNSTGTVSPVVDIGISTKANSLLISWSHGSGNVRYRLMLMDKG 240
Db 181 IAITAVSGGKRSPSVYTNSTGTVSPVVDIGISTKANSLLISWSHGSGNVRYRLMLMDKG 240
Qy 241 ILVHGGVVDKHTSYAFHGLSPGYLYNLTVMTAAAGLQNTYRWKLVRTAPMEVSNLKYTND 300
Db 241 ILVHGGVVDKHTSYAFHGLSPGYLYNLTVMTAAAGLQNTYRWKLVRTAPMEVSNLKYTND 300
Qy 301 GSLSLKVQWQRPVGNVDSYNTILSHKGTIKESRVLAPMITETHFKELVFGRLYQVTVSC 360
Db 301 GSLSLKVQWQRPVGNVDSYNTILSHKGTIKESRVLAPMITETHFKELVFGRLYQVTVSC 360
Qy 361 VSGELSAQKMAVGRTTPDKVANLEANNNGRMRSILVSWSPAGDWEQYRILLFNDSVLL 420
Db 361 VSGELSAQKMAVGRTTPDKVANLEANNNGRMRSILVSWSPAGDWEQYRILLFNDSVLL 420
Qy 421 NITVKGKEETOYVMDDTGLVPGROYEVEVIVESGNLKNSERCQGRTPVLAVLQLRVKHANE 480
Db 421 NITVKGKEETOYVMDDTGLVPGROYEVEVIVESGNLKNSERCQGRTPVLAVLQLRVKHANE 480
Qy 481 TSLSIMMOTPVAEWEKYIISLADRDLLIHKSLSKDAKEFTFDLPGRKYMATVTSIGS 540
Db 481 TSLSIMMOTPVAEWEKYIISLADRDLLIHKSLSKDAKEFTFDLPGRKYMATVTSIGS 540
Qy 541 DLKNSSSVKGRTPVPAQVTDLHVANQGMNTSLFTNWTQAOQGDVEFYQVLLIHENYVKNES 600
Db 541 DLKNSSSVKGRTPVPAQVTDLHVANQGMNTSLFTNWTQAOQGDVEFYQVLLIHENYVKNES 600
Qy 601 ISSETSRYSFHSLSKSGSLYSVVTTSVGGISSRQVVGRTVPSSVSGVTNNSGRNDYL 660
Db 601 ISSETSRYSFHSLSKSGSLYSVVTTSVGGISSRQVVGRTVPSSVSGVTNNSGRNDYL 660
Qy 661 SVSNLVAPEGVDVNEVTLSDHDKVQVSLVIKSVRECSFSLTPGRLYTVTITTRSGKYE 720
Db 661 SVSNLVAPEGVDVNEVTLSDHDKVQVSLVIKSVRECSFSLTPGRLYTVTITTRSGKYE 720
Qy 721 NHSFQERTVPDKVQGVSVNSARSYLRVSWHATGDFDHYEVTIKNKNFIQTSIPK 780
Db 721 NHSFQERTVPDKVQGVSVNSARSYLRVSWHATGDFDHYEVTIKNKNFIQTSIPK 780
Qy 781 SENECVFQVLPGRLYSVTVTTKSGQYEAENGRTIPEPVKDLTLRNESTEDLHVTVS 840
Db 781 SENECVFQVLPGRLYSVTVTTKSGQYEAENGRTIPEPVKDLTLRNESTEDLHVTVS 840
Qy 841 GANGDVDDQVEIQLEFNDKMFPPFHLVNTATEYRFTSLTPGROYKILVLTISGDVQOSAF 900
Db 841 GANGDVDDQVEIQLEFNDKMFPPFHLVNTATEYRFTSLTPGROYKILVLTISGDVQOSAF 900
Qy 901 IEGFTVPSAVKNIHI1SPNGATDSLTVNMTPGGDDVSYTVSAFRHSQKVSQTIPIKHVE 960
Db 901 IEGFTVPSAVKNIHI1SPNGATDSLTVNMTPGGDDVSYTVSAFRHSQKVSQTIPIKHVE 960
Qy 961 HTHFRLEAGQYQIMTASVSGSLKQINNVGRTVPASVQGVADNAYSSSLIVSWQKAA 1020
Db 961 HTHFRLEAGQYQIMTASVSGSLKQINNVGRTVPASVQGVADNAYSSSLIVSWQKAA 1020
Qy 1021 GVAERVDILLTENGILLRNTSEPAATKQHKFEDLTPGKYKIQILTVSGGLSKEAQTE 1080
Db 1021 GVAERVDILLTENGILLRNTSEPAATKQHKFEDLTPGKYKIQILTVSGGLSKEAQTE 1080
Qy 1081 GRTPVAAVTDLRTITENSTRHLSFRWTASEGELSWYNIFLYNDGNLQERQAVDPLVQSFS 1140
Db 1081 GRTPVAAVTDLRTITENSTRHLSFRWTASEGELSWYNIFLYNDGNLQERQAVDPLVQSFS 1140
Qy 1141 FQNLQGRMYKMVIIVTHSGBELSNESFIFGRTPVSPASVSHLRGSNNRNTTDSLWFNWSPASGD 1200

```

Db 1141 FQNLQGRMYRWIVTHSGELSNESFIFGRVTPASVSHLRGSRNRTTDSLWFWNSPASGD 1200  
Qy 1201 FDFVELILYNPGTKKENWKKDKLTETWRFQGLVPRKYVLVWVTHSGDLSNKTAEERSTA 1260  
Db 1201 FDFVELILYNPGTKKENWKKDKLTETWRFQGLVPRKYVLVWVTHSGDLSNKTAEERSTA 1260  
Qy 1261 PSPSLMSFADIANTSLAITWKGPPDWDYNDFFELQWLPDALTVPFNPNRKRSEGRIVY 1320  
Db 1261 PSPSLMSFADIANTSLAITWKGPPDWDYNDFFELQWLPDALTVPFNPNRKRSEGRIVY 1320  
Qy 1321 GLRPGSYQFNVTSGSWKTYSKPIFGSVRTKPKDKI QNLHCRPQNSTAIACSWIPDPS 1380  
Db 1321 GLRPGSYQFNVTSGSWKTYSKPIFGSVRTKPKDKI QNLHCRPQNSTAIACSWIPDPS 1380  
Qy 1381 DFDGYSIECRKMDTQEVFEFSRLEKESLLNIMLVPHKRYLSIKVQSAGMTSEVVEDS 1440  
Db 1381 DFDGYSIECRKMDTQEVFEFSRLEKESLLNIMLVPHKRYLSIKVQSAGMTSEVVEDS 1440  
Qy 1441 TITMIDRPPPPPHIRVNEKDVLSKSGSINFVNCSWFSDTNGAVKYFTVVVREADGSDE 1500  
Db 1441 TITMIDRPPPPPHIRVNEKDVLSKSGSINFVNCSWFSDTNGAVKYFTVVVREADGSDE 1500  
Qy 1501 LKPEQHQPLPSLEYRHNASIRVYOTNYPFASKCAENPNNSKSFNIIKLGAEWESLGKKRD 1560  
Db 1501 LKPEQHQPLPSLEYRHNASIRVYOTNYPFASKCAENPNNSKSFNIIKLGAEWESLGKKRD 1560  
Qy 1561 PTQOKFCGDLKPHATYRISIRAFATQLFDEDLKEFTKPLYSDFPSLPITTESEPLFGAI 1620  
Db 1561 PTQOKFCGDLKPHATYRISIRAFATQLFDEDLKEFTKPLYSDFPSLPITTESEPLFGAI 1620  
Qy 1621 EGVSAGLFLIGMLVAVVALLICROKVSNGRERPARLSIRDRPLSVHLNLGQGNKRTS 1680  
Db 1621 EGVSAGLFLIGMLVAVVALLICROKVSNGRERPARLSIRDRPLSVHLNLGQGNKRTS 1680  
Qy 1681 CPIKINQFEGHFMKLOADSNYLLSKEYEELKOVGRNQSCDIALLPENRGNRYNNILPYD 1740  
Db 1681 CPIKINQFEGHFMKLOADSNYLLSKEYEELKOVGRNQSCDIALLPENRGNRYNNILPYD 1740  
Qy 1741 ATRVKLSNVDDDCSDYINASYIPGNPNFRREYIVTQGLPCTKDDFWKWVWEQNVHIVM 1800  
Db 1741 ATRVKLSNVDDDCSDYINASYIPGNPNFRREYIVTQGLPCTKDDFWKWVWEQNVHIVM 1800  
Qy 1801 VTQCVKGRVKCDHYWADODSLYYGDLIIQLMSESVLPWTIREFKICGEBEOLDARLI 1860  
Db 1801 VTQCVKGRVKCDHYWADODSLYYGDLIIQLMSESVLPWTIREFKICGEBEOLDARLI 1860  
Qy 1861 RHEHYTWPDHGVPETTQSLIQFVRTVDRDYINRSPGAGPTVHCSAGVGRGTGFIALDRI 1920  
Db 1861 RHEHYTWPDHGVPETTQSLIQFVRTVDRDYINRSPGAGPTVHCSAGVGRGTGFIALDRI 1920  
Qy 1921 LQQLDSKSDVDIYGAVHDLRLHRVHMVQTECQVYVHLHQCVRDVLRLARKLRSEQNPLFPI 1980  
Db 1921 LQQLDSKSDVDIYGAVHDLRLHRVHMVQTECQVYVHLHQCVRDVLRLARKLRSEQNPLFPI 1980  
Qy 1981 YENVNPEYHRDPVYSRH 1997  
Db 1981 YENVNPEYHRDPVYSRH 1997

## RESULT 3

AAE20278

ID AAE20278 standard; protein; 1997 AA.

XX AC AAE20278;

XX DT 18-JUN-2002 (first entry)

XX DE Human lung specific gene (LSG) protein #16.

XX KW Human; lung specific gene; LSG; lung embryonic development; cytostatic;

XX KW lung cancer; vaccine; gene therapy; non-cancerous lung disease; tumour.

XX

OS Homo sapiens.

XX Key Location/Qualifiers  
FH Region 145..155  
FT /note= "Antigenic epitope"  
FT Region 453..465  
FT /note= "Antigenic epitope"  
FT Region 650..660  
FT /note= "Antigenic epitope"  
FT Region 714..735  
FT /note= "Antigenic epitope"  
FT Region 802..823  
FT /note= "Antigenic epitope"  
FT Region 1040..1062  
FT /note= "Antigenic epitope"  
FT Region 1121..1132  
FT /note= "Antigenic epitope"  
FT Region 1281..1295  
FT /note= "Antigenic epitope"  
FT Region 1307..1318  
FT /note= "Antigenic epitope"  
FT Region 1334..1347  
FT /note= "Antigenic epitope"  
FT Region 1351..1369  
FT /note= "Antigenic epitope"  
FT Region 1376..1408  
FT /note= "Antigenic epitope"  
FT Region 1532..1545  
FT /note= "Antigenic epitope"  
FT Region 1644..1655  
FT /note= "Antigenic epitope"  
FT Region 1671..1681  
FT /note= "Antigenic epitope"  
FT Region 1725..1738  
FT /note= "Antigenic epitope"  
FT Region 1748..1759  
FT /note= "Antigenic epitope"  
FT Region 1804..1825  
FT /note= "Antigenic epitope"  
FT Region 1885..1898  
FT /note= "Antigenic epitope"  
FT Region 1967..1976  
FT /note= "Antigenic epitope"

WO200208278-A2.

31-JAN-2002.

20-JUL-2001; 2001WO-US022949.

21-JUL-2000; 2000US-0219834P.

(DIAD-) DIADEXUS INC.

Macina RA, Nair M, Chen S;

WPI; 2002-268964/31.

Novel lung specific gene useful for identifying, diagnosing, monitoring,

staging, imaging and treating lung cancer and non-cancerous disease

states in lung, for gene therapy, and for identifying lung tissue.

Claim 2; Page 185-192; 197pp; English.

The present invention relates to lung specific genes (LSG) and their

corresponding polypeptides. LSG is useful for identifying, diagnosing,

monitoring, staging, imaging and treating lung tissue, monitoring and modifying

disease states in lung, identifying lung tissue, monitoring and modifying

lung embryonic development and differentiation, in gene therapy, as

hybridisation probes, to detect LSG mRNA as a marker for lung cancer, as

research reagents and materials for discovery of treatments and

diagnostics to human disease, to detect complementary polynucleotides,

and for chromosome identification. An antibody which binds LSG is useful

CC



CC to detect or image localisation of LSG in a patient for detecting or  
CC diagnosing a disease or condition, for preventing the onset and treatment  
CC of lung cancer, to isolate or to identify clones expressing LSG  
CC polypeptides, to purify LSG polypeptides, and to target tumours  
CC expressing LSG. The present sequence is human LSG protein  
XX  
SQ Sequence 1997 AA;

Query Match 99.8%; Score 10460; DB 5; Length 1997;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1993; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLSHGAGLAWITLSLQGLAEPCNFTLAESKASHSHSVSIQWRIILGSPCNFSLIYSS 60  
DB 1 MLSHGAGLAWITLSLQGLAEPCNFTLAESKASHSHSVSIQWRIILGSPCNFSLIYSS 60

QY 61 DTLGAALCPTFRIDNTTYGNCNLQDLQAGTIYFNKIIISLDEERTVVLQDPLPPARFGVSK 120  
DB 61 DTLGAALCPTFRIDNTTYGNCNLQDLQAGTIYFNKIIISLDEERTVVLQDPLPPARFGVSK 120

QY 121 EKTTSGLHVVWTPSSGKVTSEYVOLFDENNQKIQGVQIQESTSWNEYTFPNLTAGSKYN 180  
DB 121 EKTTSGLHVVWTPSSGKVTSEYVOLFDENNQKIQGVQIQESTSWNEYTFPNLTAGSKYN 180

QY 181 IAITAVSGGKESFVYTNSTGTPSPVKDIDIGISTKANSLIISWSHSGNVERYRLMLMDKG 240  
DB 181 IAITAVSGGKESFVYTNSTGTPSPVKDIDIGISTKANSLIISWSHSGNVERYRLMLMDKG 240

QY 241 ILVHGGVVDKATSAFHLGSPGLYXNLVTMTAEAGLQNYRWKLVRTAPMEVSNLKVTDND 300  
DB 241 ILVHGGVVDKATSAFHLGSPGLYXNLVTMTAEAGLQNYRWKLVRTAPMEVSNLKVTDND 300

QY 301 GSLTSLKVKQRPNGVDSYNITLSHKGTIKESRVLA PWITETHFKELVPGRLYQVTVSC 360  
DB 301 GSLTSLKVKQRPNGVDSYNITLSHKGTIKESRVLA PWITETHFKELVPGRLYQVTVSC 360

QY 361 VSGELSAQKMAVGRTPFDKVALEANNNGRSLVSVSPAGDWEQYRILLFNDVSVLL 420  
DB 361 VSGELSAQKMAVGRTPFDKVALEANNNGRSLVSVSPAGDWEQYRILLFNDVSVLL 420

QY 421 NITVGKEETQVMDTGLVPGQYQVEVEVIVESGNLKNSEKQCGRTVPLAVLQLRKXANE 480  
DB 421 NITVGKEETQVMDTGLVPGQYQVEVEVIVESGNLKNSEKQCGRTVPLAVLQLRKXANE 480

QY 481 TSLSIMMOTPVAEWEKYIISLADRDLLIHKSLSKDAKEFTFDLPGRKYMATVTSISG 540  
DB 481 TSLSIMMOTPVAEWEKYIISLADRDLLIHKSLSKDAKEFTFDLPGRKYMATVTSISG 540

QY 541 DLKNSSVKGRTPVPAQVTDLHVANQGMTSSLTFTWTOAGDVEFYQVLLIHENYVVKNES 600  
DB 541 DLKNSSVKGRTPVPAQVTDLHVANQGMTSSLTFTWTOAGDVEFYQVLLIHENYVVKNES 600

QY 601 ISSETSRYSFHLKSGSLYSVVVTVSSGSISSROVVVEGRTVPSSVSGVTNNSGRNDYL 660  
DB 601 ISSETSRYSFHLKSGSLYSVVVTVSSGSISSROVVVEGRTVPSSVSGVTNNSGRNDYL 660

QY 661 SVSMLVAPGDVDNTEVTLSDHDKGVQSVLIIVAKSVRECSFSLTPGRLYTVTITRSKYE 720  
DB 661 SVSMLVAPGDVDNTEVTLSDHDKGVQSVLIIVAKSVRECSFSLTPGRLYTVTITRSKYE 720

QY 721 NHPFSQERTVPDKVQGVSVNSASDYLVRVSWVHATGDFDHYEYTIKNNFIOTKSIPIK 780  
DB 721 NHPFSQERTVPDKVQGVSVNSASDYLVRVSWVHATGDFDHYEYTIKNNFIOTKSIPIK 780

QY 781 SENECVFQVLPVGRLYSVTVTKSGOYEANQNGRTIPEPVKDLTLNRSTEDLHVTWS 840  
DB 781 SENECVFQVLPVGRLYSVTVTKSGOYEANQNGRTIPEPVKDLTLNRSTEDLHVTWS 840

QY 841 GANGDVQDYEIQLLFNDMKVFPFPHLVNTATEYFTSLTPGRQYKILVLTISGDVQOSAF 900  
DB 841 GANGDVQDYEIQLLFNDMKVFPFPHLVNTATEYFTSLTPGRQYKILVLTISGDVQOSAF 900

QY 901 IEGFTVPSAVKNIHISPNGATDSLTVNWTGGGDVDSYTVSAFRHSQKVSQTIPIKHVFE 960

DB 901 IEGFTVPSAVKNIHISPNGATDSLTVNWTGGGDVDSYTVSAFRHSQKVSQTIPIKHVFE 960

QY 961 HTFHLRLEAGQYQIMIASVSGSLKNQINNVGRTVPASVQGVQVIADNAYSSVLIIVSWQKAA 1020  
DB 961 HTFHLRLEAGQYQIMIASVSGSLKNQINNVGRTVPASVQGVQVIADNAYSSVLIIVSWQKAA 1020

QY 1021 GVAERYDILALTENGILLRNTSEBATTQKHFEDELTPGKKYKIQILTVSGGLFSKEAQTE 1080  
DB 1021 GVAERYDILALTENGILLRNTSEBATTQKHFEDELTPGKKYKIQILTVSGGLFSKEAQTE 1080

QY 1081 GRTVPAAVTDLRITENSTRHLSFRWTASEGELSWYNIFLYNPDGNLQERAOVDPLVQSFS 1140  
DB 1081 GRTVPAAVTDLRITENSTRHLSFRWTASEGELSWYNIFLYNPDGNLQERAOVDPLVQSFS 1140

QY 1141 FQNLQGRMYKMWIVTHSGELSNESFIFGRTPVASVSHLRGSNRRNTTDSLWFNWSPASGD 1200  
DB 1141 FQNLQGRMYKMWIVTHSGELSNESFIFGRTPVASVSHLRGSNRRNTTDSLWFNWSPASGD 1200

QY 1201 FDFVELILYNPNTGKKNWKDKOLTEWRFOGLVPGRKVYLWVTHSGDLSNKVTAESRTA 1260  
DB 1201 FDFVELILYNPNTGKKNWKDKOLTEWRFOGLVPGRKVYLWVTHSGDLSNKVTAESRTA 1260

QY 1261 PSPPLMSFADIANSLAITWKGPDMTDYNDPELOMLPRDALTVFNPYNNRKSEGRIVY 1320  
DB 1261 PSPPLMSFADIANSLAITWKGPDMTDYNDPELOMLPRDALTVFNPYNNRKSEGRIVY 1320

QY 1321 GLRPGRSYQFNVTKVSQDSWKTSKPIFGSVRTKPKDKIQNLHCRPQNSTAIACSWIPDDS 1380  
DB 1321 GLRPGRSYQFNVTKVSQDSWKTSKPIFGSVRTKPKDKIQNLHCRPQNSTAIACSWIPDDS 1380

QY 1381 DFDGYSIECKRMDTOEVEFSRKLKESKSLINIMLVPHKRYLVSIVKQSGAMTSEVVEDS 1440  
DB 1381 DFDGYSIECKRMDTOEVEFSRKLKESKSLINIMLVPHKRYLVSIVKQSGAMTSEVVEDS 1440

QY 1441 TITWIDRPPPPPHIRVNEKDVLSKSSINFTVNCNMFSDTNGAVKFTVVVRADSGDE 1500  
DB 1441 TITWIDRPPPPPHIRVNEKDVLSKSSINFTVNCNMFSDTNGAVKFTVVVRADSGDE 1500

QY 1501 LKPEQQHPLPSYLEYRHNASIRVYQTNVYFASKCAENPNSNSKSFNIIKLGAEMLSGGKCD 1560  
DB 1501 LKPEQQHPLPSYLEYRHNASIRVYQTNVYFASKCAENPNSNSKSFNIIKLGAEMLSGGKCD 1560

QY 1561 PTQOKFCGDLKPHHTAYRISIRAPFTQLPDEDLKFTKPLYSDDTFFSLPITTESPLFGAI 1620  
DB 1561 PTQOKFCGDLKPHHTAYRISIRAPFTQLPDEDLKFTKPLYSDDTFFSLPITTESPLFGAI 1620

QY 1621 EGVSAGLFLIGMLVAVVALLICRQKVSCHGRERPSARLSIRDRPLSVHLNIGQKGNRKT 1680  
DB 1621 EGVSAGLFLIGMLVAVVALLICRQKVSCHGRERPSARLSIRDRPLSVHLNIGQKGNRKT 1680

QY 1681 CPKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQCDAIALLPENRGKNNYNNILPYD 1740  
DB 1681 CPKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQCDAIALLPENRGKNNYNNILPYD 1740

QY 1741 ATRVKLSNVDDDDPCSDVINASYIPGNFRREYIVTQGPLPGTKDDDFWQVWQVNHVIM 1800  
DB 1741 ATRVKLSNVDDDDPCSDVINASYIPGNFRREYIVTQGPLPGTKDDDFWQVWQVNHVIM 1800

QY 1801 VTQCVKGRVKCDHYWPAQDODSLYGGDLILQMLSESVLPBWTIREFKICGEEQLDAHRLI 1860  
DB 1801 VTQCVKGRVKCDHYWPAQDODSLYGGDLILQMLSESVLPBWTIREFKICGEEQLDAHRLI 1860

QY 1861 RHPHYTVPDHPGPEPTTQSLIQFVRTVDRVYNRSPGAGPTTVHCSAGVGRGTGTIALDRI 1920  
DB 1861 RHPHYTVPDHPGPEPTTQSLIQFVRTVDRVYNRSPGAGPTTVHCSAGVGRGTGTIALDRI 1920

QY 1921 LQQLDSKDSVDIYGAVHDLRLHRVHMVQTECQVYVYLHCQVRDVLRAKRLSEQNPLFPI 1980  
DB 1921 LQQLDSKDSVDIYGAVHDLRLHRVHMVQTECQVYVYLHCQVRDVLRAKRLSEQNPLFPI 1980

QY 1981 YENVNPEYHDDPVYSRH 1997

Db	1981 YENVNPEYHRDPVYSRH 1997	CC	PTP10D, Tec, or EDRP gene product. The present sequence represents human
RESULT 4		CC	• protein tyrosine phosphatase receptor type B precursor (PTPRB), which is
ABR57179		CC	a human PTP10B homologous sequence. Human PTPRB is located to chromosome
ID	ABR57179 standard; protein; 1997 AA.	CC	12
XX		XX	
AC	ABR57179;	XX	Sequence 1997 AA;
DT	02-SEP-2003 (first entry)		
XX			
DE	Human PTPRB protein SEQ ID NO:4.		
XX			
KW	Tec; protein tyrosine kinase; protein tyrosine phosphatase; PTP10D;		
KW	egg derived tyrosine phosphatase; EDRP; antidiabetic; hypotensive;		
KW	cardiant; antilipaeamic; osteopathic; cytosstatic; anorectic; obesity;		
KW	immunomodulator; gene therapy; metabolic disease; eating disorder;		
KW	body weight regulation disorder; cachexia; diabetes mellitus; cancer;		
KW	hypertension; coronary heart disease; hypercholesterolaemia; gallstone;		
KW	dyslipidaemia; osteoarthritis; sleep apnea; human; chromosome 12;		
KW	protein tyrosine phosphatase receptor type B precursor; PTPRB.		
OS	Homo sapiens.		
FN	WO2003047611-A2.		
XX			
PD	12-JUN-2003.		
XX			
PF	04-DEC-2002; 2002WO-EP013744.		
XX			
PR	04-DEC-2001; 2001EP-00128844.		
PR	07-DEC-2001; 2001EP-00129138.		
PR	02-JAN-2002; 2002EP-00000010.		
XX			
PA	(DEVE-) DEVELOPEN ENTWICKLUNGSBIOLOGISCHE FORSCH.		
XX			
PI	Meise M, Eulenberger K, Fritsch R, Haeder T, Broenner G;		
PI	Steuernagel A;		
PI			
DR	WPI; 2003-532801/50.		
DR	N-PSDB; ACC79776.		
XX			
PT	New compositions comprising tyrosine phosphatase PTP10D, protein tyrosine		
PT	kinase Tec or egg-derived tyrosine phosphatase genes or proteins, useful		
PT	for treating or preventing metabolic diseases, e.g. as obesity or		
PT	cachexia.		
XX			
PS	Claim 2; Fig 8B; 83pp; English.		
XX			
CC	The present invention describes a pharmaceutical composition comprising a		
CC	nucleic acid (I) protein tyrosine phosphatase PTP10D, non-receptor		
CC	protein tyrosine kinase Tec, egg derived tyrosine phosphatase (EDRP) gene		
CC	family or encoded polypeptide, fragment or variant of nucleic acid		
CC	molecule or polypeptide, an antibody, an aptamer or receptor recognising		
CC	a nucleic acid molecule of PTP10D, Tec, or EDRP gene family or encoded		
CC	polypeptide, and a carrier, diluent and/or adjuvant. The pharmaceutical		
CC	composition can have antidiabetic, hypotensive, cardiant, antilipaeamic,		
CC	osteopathic, cytosstatic, anorectic and immunomodulator activities, and		
CC	can be used in gene therapy. The composition is useful for the		
CC	manufacture of an agent for detecting and/or verifying, for treating and		
CC	alleviating and/or preventing a disorder, including metabolic diseases		
CC	such as obesity and other body weight regulation disorders, as well as		
CC	related disorders such as eating disorder, cachexia, diabetes mellitus,		
CC	hypertension, coronary heart disease, hypercholesterolaemia, of the		
CC	dyslipidaemia, osteoarthritis, gallstones, cancers (cancers of the		
CC	reproductive organ), sleep apnea, and other diseases, in cells, cell		
CC	masses, organs and/or subjects. The components of the composition may		
CC	also be used in controlling the function of a gene and/or gene product		
CC	which is influenced and/or modified by a PTP10D, Tec, or EDRP homologous		
CC	polypeptide, and for identifying substances capable of interacting with a		
CC	PTP10B, Tec or EDRP homologous polypeptide. The nucleic acid molecule of		
CC	PTP10D, Tec, or EDRP family or their fragments, may be used in the		
CC	preparation of a non-human animal which over- or under-expresses the		

Query Match	99.8%;	Score	10460;	DB	7;	Length	1997;
Best Local Similarity	99.7%;	Pred. No.	0;				
Matches	1992;	Conservative	3;	Mismatches	2;	Indels	0;
Gaps	0;						
Qy	1	MLSHGAGLALWITLSLLQTGLAEPFCNFTLASKASSHSVSITOWRILGSPCNFSLIYSS	60				
Db	1	MLSHGAGLALWITLSLLQTGLAEPFCNFTLASKASSHSVSITOWRILGSPCNFSLIYSS	60				
Qy	61	DTLGAALCPTFRIDNTTYGNCLODQAGTIYFKIISLDEERTVVLQTDPLPARGVSK	120				
Db	61	DTLGAALCPTFRIDNTTYGNCLODQAGTIYFKIISLDEERTVVLQTDPLPARGVSK	120				
Qy	121	EKTTSTGLHVMWTPSSGKVTSEYVQLFDENNQIKGVQIQESTSWNEYTFNLTAGSKYN	180				
Db	121	EKTTSTGLHVMWTPSSGKVTSEYVQLFDENNQIKGVQIQESTSWNEYTFNLTAGSKYN	180				
Qy	181	IAITAVSGGKRSPSVYTNSTGTPSPVKDIGISTKANSLLISWSHSGSNVERYLMLMDKG	240				
Db	181	IAITAVSGGKRSPSVYTNSTGTPSPVKDIGISTKANSLLISWSHSGSNVERYLMLMDKG	240				
Qy	241	ILVHGGVVDKHTSYAFHGLSPGYLYNLVTMTAAGLQNYRWKLVRTAPMEVSNLKVTDND	300				
Db	241	ILVHGGVVDKHTSYAFHGLTPGYLYNLVTMTAAGLQNYRWKLVRTAPMEVSNLKVTDND	300				
Qy	301	GSUTSLKVKWQRPNGVDSYNTLSHGKTIKESRVLAPWITETHFKELVPGRLYQVTVSC	360				
Db	301	GSUTSLKVKWQRPNGVDSYNTLSHGKTIKESRVLAPWITETHFKELVPGRLYQVTVSC	360				
Qy	361	VSGELSAQKMAVGRTFPDKVANLEANNNGRMRLSVVSWSPGADWQYRILLFNDSVLL	420				
Db	361	VSGELSAQKMAVGRTFPDKVANLEANNNGRMRLSVVSWSPGADWQYRILLFNDSVLL	420				
Qy	421	NITVKEETQYVMDTGLVPGROYEVEVIVESGNLKNSERCQGRTPVLAQLRVKHANE	480				
Db	421	NITVKEETQYVMDTGLVPGROYEVEVIVESGNLKNSERCQGRTPVLAQLRVKHANE	480				
Qy	481	TSLSIMWQTPVAEWKYYIISLADRLILLHKSLSKDAKFTFTDLVPGRYKMATVTSIG	540				
Db	481	TSLSIMWQTPVAEWKYYIISLADRLILLHKSLSKDAKFTFTDLVPGRYKMATVTSIG	540				
Qy	541	DLKNSSSVKGRTVPAQVTDLHVANQWTSLSLFTNWTQAQGDVEFYQVLLIHENNVKNES	600				
Db	541	DLKNSSSVKGRTVPAQVTDLHVANQWTSLSLFTNWTQAQGDVEFYQVLLIHENNVKNES	600				
Qy	601	ISSETSRYSFHSLSKSLYSVWVTVVSGGSSRQVVEGRTVPSSVSGVTNNSGRNDYL	660				
Db	601	ISSETSRYSFHSLSKSLYSVWVTVVSGGSSRQVVEGRTVPSSVSGVTNNSGRNDYL	660				
Qy	661	SVSWLAPGDVDNNEYVTLSDHGKVVQSLVIAKSVRECSFSLTPGRLYTVTTITRSKYE	720				
Db	661	SVSWLAPGDVDNNEYVTLSDHGKVVQSLVIAKSVRECSFSLTPGRLYTVTTITRSKYE	720				
Qy	721	NHSFSQERTVPDKVQGVSVNSARSYLRVSWVHATGDFDHYEVTIKNNKNTFTQTSIPK	780				
Db	721	NHSFSQERTVPDKVQGVSVNSARSYLRVSWVHATGDFDHYEVTIKNNKNTFTQTSIPK	780				
Qy	781	SENECVFQVLPGRLYSVTVTTKSGQYEAQNGRTIPEPVKDLTLRNESTDLHVTVWS	840				
Db	781	SENECVFQVLPGRLYSVTVTTKSGQYEAQNGRTIPEPVKDLTLRNESTDLHVTVWS	840				
Qy	841	GANGDVQDQEIQLLFNDKMFVFPFHLVNTATEYRFTSLTPGROYKILVLTISGDVQOSAF	900				
Db	841	GANGDVQDQEIQLLFNDKMFVFPFHLVNTATEYRFTSLTPGROYKILVLTISGDVQOSAF	900				
Qy	901	IEGFTVPSAVKNTHISPNAGATSLTVNWTGGGDVDSYTVSAFPHSKQVDSQTIPIKHVPE	960				
Db	901	IEGFTVPSAVKNTHISPNAGATSLTVNWTGGGDVDSYTVSAFPHSKQVDSQTIPIKHVPE	960				

Db	901	LEGFTVPSAVKNIHISPNATDLSLTWNVTGCGDVSYSYTVSAPRHSQKVDSTQIPKHVFE	960
Qy	961	HTFHRLEAGEQYQIMIASVSGSLKNQINNVGRTPVPSVQGVIAADNAYSSYSLIVSWOKAA	1020
Db	961	HTFHRLEAGEQYQIMIASVSGSLKNQINNVGRTPVPSVQGVIAADNAYSSYSLIVSWOKAA	1020
Qy	1021	GVAERYDILLTENGILLRNTSEPAATKQHKFEDLTGKKYKIOILTIVSGGLFSGKEAQTE	1080
Db	1021	GVAERYDILLTENGILLRNTSEPAATKQHKFEDLTGKKYKIOILTIVSGGLFSGKEAQTE	1080
Qy	1081	GRTVPAAVTDLRITENSTRHLSFRWTASEGELSWYNIPLYPDGNLQERAOVDPLVQSFS	1140
Db	1081	GRTVPAAVTDLRITENSTRHLSFRWTASEGELSWYNIPLYPDGNLQERAOVDPLVQSFS	1140
Qy	1141	FQNLQGRMYQWIVTHSGELSNESFIPGRTPVPSVSHLRGSRNTTDSLFWNSPASGD	1200
Db	1141	FQNLQGRMYQWIVTHSGELSNESFIPGRTPVPSVSHLRGSRNTTDSLFWNSPASGD	1200
Qy	1201	FDYVELILYNGTKENWKDKDTEWRFQGLVGRKYVWVTHSGDLSNKNVTAESRTA	1260
Db	1201	FDYVELILYNGTKENWKDKDTEWRFQGLVGRKYVWVTHSGDLSNKNVTAESRTA	1260
Qy	1261	PSPPSLMSFADIANSTLAIWKGPDPDWDYNDFFELQMLPRDALTVFNPYNNRKSSEGRIVY	1320
Db	1261	PSPPSLMSFADIANSTLAIWKGPDPDWDYNDFFELQMLPRDALTVFNPYNNRKSSEGRIVY	1320
Qy	1321	GLRPGRSYQFNKTVSGDSWTKYKPIGFSVRTKPKDKIQLNLCRPNQSTAIACSWIPDPS	1380
Db	1321	GLRPGRSYQFNKTVSGDSWTKYKPIGFSVRTKPKDKIQLNLCRPNQSTAIACSWIPDPS	1380
Qy	1381	DFDGYSTECRQMDQTEVEFSRKLKESKLLNIMMLVPHKRYLVSIKVSAGMTSEVVEDS	1440
Db	1381	DFDGYSTECRQMDQTEVEFSRKLKESKLLNIMMLVPHKRYLVSIKVSAGMTSEVVEDS	1440
Qy	1441	TITMIDRPPPPPHIRVNEKDVLSKSSINFTVNCSPFSDTNGAVKFTVVVRADGSDE	1500
Db	1441	TITMIDRPPPPPHIRVNEKDVLSKSSINFTVNCSPFSDTNGAVKFTVVVRADGSDE	1500
Qy	1501	LKPEQHQHPLPSLYRHNASIRVYQTNFYASKCAENPNNSKSFNKLGAEMESLGGKRD	1560
Db	1501	LKPEQHQHPLPSLYRHNASIRVYQTNFYASKCAENPNNSKSFNKLGAEMESLGGKRD	1560
Qy	1561	PTQKFCDCGPKPHATYRISIRATQLPDEDLKEFTKPLYSDDTFFSLPITTESPLFGAI	1620
Db	1561	PTQKFCDCGPKPHATYRISIRATQLPDEDLKEFTKPLYSDDTFFSLPITTESPLFGAI	1620
Qy	1621	EGVSAGLFLICMLVAVALLICROKVSGRERPSARLSIRDRPLSVHLNLGQGNKRTS	1680
Db	1621	EGVSAGLFLICMLVAVALLICROKVSGRERPSARLSIRDRPLSVHLNLGQGNKRTS	1680
Qy	1681	CPIKINQPEGHFMKLOADSNYLLSKEYBELKDVGRNQSCDIALLPENRGNRYNNILPYD	1740
Db	1681	CPIKINQPEGHFMKLOADSNYLLSKEYBELKDVGRNQSCDIALLPENRGNRYNNILPYD	1740
Qy	1741	ATRVKLSNVDDPCSDYINASYIPGNPRRYIVITQGPLPOTKDDFWKQWQNVHNIIVM	1800
Db	1741	ATRVKLSNVDDPCSDYINASYIPGNPRRYIVITQGPLPOTKDDFWKQWQNVHNIIVM	1800
Qy	1801	VTQCEKGRVKCDHYWPAQDQSLYYGDLILQMLSESVLPEWTIREFKICGEEQDHAHLI	1860
Db	1801	VTQCEKGRVKCDHYWPAQDQSLYYGDLILQMLSESVLPEWTIREFKICGEEQDHAHLI	1860
Qy	1861	RHFHYTVMPPDGVPTTQSLIQFVRTVRDYINRSPGAGPTVHCSAGVGRGTGTIALDRI	1920
Db	1861	RHFHYTVMPPDGVPTTQSLIQFVRTVRDYINRSPGAGPTVHCSAGVGRGTGTIALDRI	1920
Qy	1921	LQQLDSKDSVDIYGAVHDLRLHRVHMVQTECOYVYLHCQVVDVLRLARKLSEQNPLFPI	1980
Db	1921	LQQLDSKDSVDIYGAVHDLRLHRVHMVQTECOYVYLHCQVVDVLRLARKLSEQNPLFPI	1980
Qy	1981	YENVNPEYHRDPVYSRH	1997
Db	1981	YENVNPEYHRDPVYSRH	1997

RESULT 5  
 ID ADJ70329 standard; protein; 1997 AA.  
 XX ADJ70329;  
 AC ADJ70329;  
 XX 06-MAY-2004 (first entry)  
 DT Human heat mitochondrial protein as a therapeutic target SeqID2135.  
 XX  
 DE mitochondrial; human; screening assay; diabetes mellitus;  
 KW Huntington's disease; osteoarthritis; LHON;  
 KW Leber's hereditary optic neuropathy; LHON;  
 KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;  
 KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;  
 KW neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;  
 KW osteopathic; ophthalmological; cytostatic.  
 XX Homo sapiens.  
 OS  
 XX WO2003087768-A2.  
 XX 23-OCT-2003.  
 XX 04-APR-2003; 2003WO-US010870.  
 XX 12-APR-2002; 2002US-0372843P.  
 XX 17-JUN-2002; 2002US-0389987P.  
 XX 20-SEP-2002; 2002US-0412418P.  
 XX  
 XX (MITO-) MITOKOR.  
 XX (BUCK-) BUCK INST AGE RES.  
 XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GW;  
 PI Warnock DE;  
 XX WPI; 2003-845369/78.  
 DR  
 XX  
 XX Identifying a mitochondrial target for drug screening assays and for  
 treating diseases associated with altered mitochondrial function,  
 comprises detecting a modified polypeptide in a sample and correlating  
 with the disease.  
 XX  
 XX Claim 1; SEQ ID NO 2135; 180pp; English.  
 PS  
 XX This invention relates to novel mitochondrial targets that can be used  
 for therapeutic intervention in treating a disease associated with  
 altered mitochondrial function. Specifically, it refers to a method for  
 identifying proteins of the human heart mitochondrial proteome that are  
 useful for drug screening assays, as well as therapeutic targets. The  
 present invention describes a method for identifying such proteins that  
 can be used in the treatment of various diseases associated with altered  
 mitochondrial function including diabetes mellitus, Huntington's disease,  
 osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial  
 encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy  
 ragged red fibre syndrome (MERRF) or cancer. Accordingly, these  
 compositions have neuroprotective, nontropic, antidiabetic,  
 anticonvulsant, antiarthritic, osteopathic, ophthalmological and  
 cytostatic activities. This polypeptide sequence is a human heart  
 mitochondrial protein of the invention.  
 XX  
 XX Sequence 1997 AA;

Query Match 99.8%; Score 10460; DB 7; Length 1997;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 1992; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 MLSHGAGLAWITLSLQTLGAEPCNFTLAESKASHSVSIQWRILGSPCNFLSYSS 60  
 Db 1 MLSHGAGLAWITLSLQTLGAEPCNFTLAESKASHSVSIQWRILGSPCNFLSYSS 60

Qy 61 DTLGAALCPTFRIDNTTYGCMQLDLQAGTIYNFKIISLDEBRTVVLTDPPLPPARFGVSK 120  
Db 61 DTLGAALCPTFRIDNTTYGCMQLDLQAGTIYNFKIISLDEBRTVVLTDPPLPPARFGVSK 120  
Qy 121 EKTTSGLHVMWTPSSGKVTSEYVQLFDENNOKIQGVQIQESTSWNEYTFNLTAGSKYN 180  
Db 121 EKTTSGLHVMWTPSSGKVTSEYVQLFDENNOKIQGVQIQESTSWNEYTFNLTAGSKYN 180  
Qy 181 IAITAVSGGKSFVYTNGSTVPVKDIGISTKANSLIISWSHSGGNVERYRLMLMDKG 240  
Db 181 IAITAVSGGKSFVYTNGSTVPVKDIGISTKANSLIISWSHSGGNVERYRLMLMDKG 240  
Qy 241 ILVHGGVVDKATSYAFHGLSPGYLYNLVTVMTEAAGLQNYRWKLVRTAPMEVSNLKVTDND 300  
Db 241 ILVHGGVVDKATSYAFHGLTPGYLYNLVTVMTEAAGLQNYRWKLVRTAPMEVSNLKVTDND 300  
Qy 301 GSLTSLVKVQRPQGNVDVSNITLSHKGTIKESRVLA PWITETHFKELVPGRLYQVTVSC 360  
Db 301 GSLTSLVKVQRPQGNVDVSNITLSHKGTIKESRVLA PWITETHFKELVPGRLYQVTVSC 360  
Qy 361 VSGELSAQKMAVGRTFPDKVANLEANNRMRSLVWSPPAGDWEQVYRILLFNDVSVLL 420  
Db 361 VSGELSAQKMAVGRTFPDKVANLEANNRMRSLVWSPPAGDWEQVYRILLFNDVSVLL 420  
Qy 421 NITVKGEEQYVMDDTGLVPGRQYEVEVIVESGNLKNSERCQGRTPVLA VLQLRVKHANE 480  
Db 421 NITVKGEEQYVMDDTGLVPGRQYEVEVIVESGNLKNSERCQGRTPVLA VLQLRVKHANE 480  
Qy 481 TSLSIMQTPVAEWEKIIISLADRDLLIHKSLSKDAKEFTFDLPVGRKYMATVTSISG 540  
Db 481 TSLSIMQTPVAEWEKIIISLADRDLLIHKSLSKDAKEFTFDLPVGRKYMATVTSISG 540  
Qy 541 DLKNSSVKGRTPVAQVTDLVANQGMTSSLFTNWTQAQGDVRFYQVLLIHENVVIKNES 600  
Db 541 DLKNSSVKGRTPVAQVTDLVANQGMTSSLFTNWTQAQGDVRFYQVLLIHENVVIKNES 600  
Qy 601 ISSETRSYFSLKSGSLYSVVVTVSGGISRQVVEGRTVPSSVSGVTVNNSGRNDYL 660  
Db 601 ISSETRSYFSLKSGSLYSVVVTVSGGISRQVVEGRTVPSSVSGVTVNNSGRNDYL 660  
Qy 661 SVSHLVARGVDNVEVTLSDHGKVVQSLVIAKSVRECSFSLTPGRLYTVTITTRSGKYE 720  
Db 661 SVSHLVARGVDNVEVTLSDHGKVVQSLVIAKSVRECSFSLTPGRLYTVTITTRSGKYE 720  
Qy 721 NHPFSQERTVPDKQGVSVNSARS DYLRSVSWHATGDFDHYEVTIKNNKFIOTKSIPIK 780  
Db 721 NHPFSQERTVPDKQGVSVNSARS DYLRSVSWHATGDFDHYEVTIKNNKFIOTKSIPIK 780  
Qy 781 SENECEVFQVLPGRLYSVTVTTKSGQYEANEQNGRTIPEPVKDLTLNRSTEDLHVTS 840  
Db 781 SENECEVFQVLPGRLYSVTVTTKSGQYEANEQNGRTIPEPVKDLTLNRSTEDLHVTS 840  
Qy 841 GANGDVQYEIQLLFNDMKVPPPHLVNTATEYFTSLTPGRQYKIILVLTISGDVQOSAF 900  
Db 841 GANGDVQYEIQLLFNDMKVPPPHLVNTATEYFTSLTPGRQYKIILVLTISGDVQOSAF 900  
Qy 901 IEGFTVPSAVKNIIHISPNAGTDSLTVNWTGGGDVDSVTSAPRHSOKVDSQITIPKHVFE 960  
Db 901 IEGFTVPSAVKNIIHISPNAGTDSLTVNWTGGGDVDSVTSAPRHSOKVDSQITIPKHVFE 960  
Qy 961 HTFHRLEAGEQYQIMIASVSGSLKNQINNVGRTVPASVQGVADNAYSSYSLIVSWQKAA 1020  
Db 961 HTFHRLEAGEQYQIMIASVSGSLKNQINNVGRTVPASVQGVADNAYSSYSLIVSWQKAA 1020  
Qy 1021 GVAERYDILLITENGILLRNTSEPAATTKQHKFEDLTPGKKYKIQIILTVSGGLFSKEAQTE 1080  
Db 1021 GVAERYDILLITENGILLRNTSEPAATTKQHKFEDLTPGKKYKIQIILTVSGGLFSKEAQTE 1080  
Qy 1081 GRTPVAAVTDLRITENSTRHLSPFWTASEGELSWYNIPLYNPDGNLQERAOQVDPVLOSFS 1140  
Db 1081 GRTPVAAVTDLRITENSTRHLSPFWTASEGELSWYNIPLYNPDGNLQERAOQVDPVLOSFS 1140  
Qy 1141 FQNLLQGRMYKMVIWTHSGELSNESFIFGRTPVASVSHLRGSNRRNTTDSLWFNWPASGD 1200

Db 1141 FQNLLQGRMYKMVIWTHSGELSNESFIFGRTPVASVSHLRGSNRRNTTDSLWFNWPASGD 1200  
Qy 1201 FDFVELLLYNPNGTKKENWKDKOLTWRFOGLVPGRYKYLWVWTHSGDLNKNVTAESRTA 1260  
Db 1201 FDFVELLLYNPNGTKKENWKDKOLTWRFOGLVPGRYKYLWVWTHSGDLNKNVTAESRTA 1260  
Qy 1261 PSPPSLMSFADIANSTSLAITWKGPPDMTDYNDPELOWLPDALTVPFNPYNNRKSSEGRIVY 1320  
Db 1261 PSPPSLMSFADIANSTSLAITWKGPPDMTDYNDPELOWLPDALTVPFNPYNNRKSSEGRIVY 1320  
Qy 1321 GLRPGRSYQFNVTKVSQDSWKTSYKPIFGSVRTKPKDKI QNLHCRPQNSTAIACSWIPDPS 1380  
Db 1321 GLRPGRSYQFNVTKVSQDSWKTSYKPIFGSVRTKPKDKI QNLHCRPQNSTAIACSWIPDPS 1380  
Qy 1381 DFDGYSIECKRMDTOEVEFSRKLKESKSLINIMLVPHKRYLSIKVQSAGMTSEVVEDS 1440  
Db 1381 DFDGYSIECKRMDTOEVEFSRKLKESKSLINIMLVPHKRYLSIKVQSAGMTSEVVEDS 1440  
Qy 1441 TITMIDRPPPPPHIRVNEKDVLSISKSSINFTVNCWFSDTNGAVKYFTVVVREADGSDE 1500  
Db 1441 TITMIDRPPPPPHIRVNEKDVLSISKSSINFTVNCWFSDTNGAVKYFTVVVREADGSDE 1500  
Qy 1501 LKPEQHQHPLPSYLEYRHNASIRVYQTNYPASKCAENPNSNSKSFNIIKLGAEMLSGKRD 1560  
Db 1501 LKPEQHQHPLPSYLEYRHNASIRVYQTNYPASKCAENPNSNSKSFNIIKLGAEMLSGKRD 1560  
Qy 1561 PTQOKFCGDLKPHATYRISIRAFPTQLDFEDLKEFTKPLYSDFSSLPITTESPLFGAI 1620  
Db 1561 PTQOKFCGDLKPHATYRISIRAFPTQLDFEDLKEFTKPLYSDFSSLPITTESPLFGAI 1620  
Qy 1621 EGVSAGLFLGLMLVAVVALLICRQVSHGRERSARLSIRDRPLSVHLNLGQGNRKT 1680  
Db 1621 EGVSAGLFLGLMLVAVVALLICRQVSHGRERSARLSIRDRPLSVHLNLGQGNRKT 1680  
Qy 1681 CPIKINQFEGHFMKLQADSNLYLSKEYEELKDVGRNQSCDIALLPENRGNKRYNNIILPYD 1740  
Db 1681 CPIKINQFEGHFMKLQADSNLYLSKEYEELKDVGRNQSCDIALLPENRGNKRYNNIILPYD 1740  
Qy 1741 ATRVKLSNVDDDDPCSDYINASYIPGNFRREYIVTOGPLPGTKDDFWKMWQGNVHNVM 1800  
Db 1741 ATRVKLSNVDDDDPCSDYINASYIPGNFRREYIVTOGPLPGTKDDFWKMWQGNVHNVM 1800  
Qy 1801 VTQCVKGRVKCDHYVPADQDSLYYGLIILQMLSESVLPETWITREFKICGEEQLDAHRLI 1860  
Db 1801 VTQCVKGRVKCDHYVPADQDSLYYGLIILQMLSESVLPETWITREFKICGEEQLDAHRLI 1860  
Qy 1861 RHFHYTVPDGHVPETTSQSLIQFVRTVRYDINRSPGAGTVPVHCSAGVGTGTGFIALDRI 1920  
Db 1861 RHFHYTVPDGHVPETTSQSLIQFVRTVRYDINRSPGAGTVPVHCSAGVGTGTGFIALDRI 1920  
Qy 1921 LQOLDSKDSVDIYGAVHDLRLHVRHMVQTECOVYLLHQCVRDVLRARKLSEQNPLFPI 1980  
Db 1921 LQOLDSKDSVDIYGAVHDLRLHVRHMVQTECOVYLLHQCVRDVLRARKLSEQNPLFPI 1980  
Qy 1981 YENVNPEYHRDPVYSRH 1997  
Db 1981 YENVNPEYHRDPVYSRH 1997

## RESULT 6

AAW78821  
ID AAW78821 standard; protein; 1997 AA.

XX AAW78821;

XX 06-NOV-2001 (first entry)

XX Human protein SEQ ID NO 1483.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW	nervous system disorder; arthritis; inflammation.	Qy	301	GSLSLKVKKWQRPNGVDSYNITLSHGKTIKESRVLAPWITETHFKELVGRLYQVTVSC	360
XX					
OS	Homo sapiens.	Db	301	GSLSLKVKKWQRPNGVDSYNITLSHGKTIKESRVLAPWITETHFKELVGRLYQVTVSC	360
XX					
PN	WO200157190-A2.	Qy	361	VSGELSAQKMAVGRTFPDKVANLEANNGRMRSILVSVSWSPAGDQWQYRILLFNDVSVLL	420
XX					
XX	09-AUG-2001.	Db	361	VSGELSAQKMAVGRTFPDKVANLEANNGRMRSILVSVSWSPAGDQWQYRILLFNDVSVLL	420
XX					
PF	05-FEB-2001; 2001WO-US004098.	Qy	421	NITVGKEETOYVMDDTGLVPGROYEVEVIVESGNLKNSERCQGRTPVPLAVLQVRKANE	480
XX					
PR	03-FEB-2000; 2000US-00496914.	Db	421	NITVGKEETOYVMDDTGLVPGROYEVEVIVESGNLKNSERCQGRTPVPLAVLQVRKANE	480
XX					
PR	27-APR-2000; 2000US-00560875.	Qy	481	TSLSIMMOTPAVEWEKVIISLADRDLLIHKSLSDAKKEFTTDLVPCRKYMTVTSISG	540
XX					
PR	20-JUN-2000; 2000US-00598075.	Db	481	TSLSIMMOTPAVEWEKVIISLADRDLLIHKSLSDAKKEFTTDLVPCRKYMTVTSISG	540
XX					
PR	19-JUL-2000; 2000US-00620325.	Qy	541	DLKNSSSVKGRTPVPAQVTDLFVANQGMTSSLTNTWTOAQGDVEFYQVLLIHENYVVKNES	600
XX					
PR	01-SEP-2000; 2000US-00654936.	Db	541	DLKNSSSVKGRTPVPAQVTDLFVANQGMTSSLTNTWTOAQGDVEFYQVLLIHENYVVKNES	600
XX					
PR	15-SEP-2000; 2000US-00663561.	Qy	601	ISSETSRYSFHLKSGSLYSVVVTTVSGGISRQVVEGRTVPSSVSGVTVNNSGRNDYL	660
XX					
PR	20-OCT-2000; 2000US-00693325.	Db	601	ISSETSRYSFHLKSGSLYSVVVTTVSGGISRQVVEGRTVPSSVSGVTVNNSGRNDYL	660
XX					
PR	30-NOV-2000; 2000US-00728422.	Qy	661	SVSWLAPGDVDNVEVTLSDHGKVVQSLVIKSVRECSFSLTGRLYTVTTITRSGKYE	720
XX					
PA	(HYSE-) HYSEQ INC.	Db	661	SVSWLAPGDVDNVEVTLSDHGKVVQSLVIKSVRECSFSLTGRLYTVTTITRSGKYE	720
XX					
XX	Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;	Qy	721	NHSFQSBQRTVPDKVQGVSVSNARSXYLRVSWHATGDFDHYEVTIKNNFIQTKSIPK	780
PI	Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;	Db	721	NHSFQSBQRTVPDKVQGVSVSNARSXYLRVSWHATGDFDHYEVTIKNNFIQTKSIPK	780
PI	Xue AJ, Yang Y, Wejhrman T, Goodrich R;	Qy	781	SENECFVQLVPGRLYSVTVTTKSGQYEAENGQNGRTIPEPVKDLTLNRSTEDLHVTWS	840
XX					
XX	WPI; 2001-476283/51.	Db	841	GANGDVQYIQLLFNDMKVFPFPHLVNTATEYRFTSLTPGRQYKILVLTISGVDQOSAF	900
DR	N-PSDB; AAK51954.	Qy	841	GANGDVQYIQLLFNDMKVFPFPHLVNTATEYRFTSLTPGRQYKILVLTISGVDQOSAF	900
XX					
XX	Nucleic acids encoding polypeptides with cytokine-like activities, useful	Qy	901	IEGFTVPSAVKNTHISPNAGTDSLTVNWTGGGDVDSYTVSAFRHSQKVDQSOTIPKHVFE	960
XX	in diagnosis and gene therapy.	Db	901	IEGFTVPSAVKNTHISPNAGTDSLTVNWTGGGDVDSYTVSAFRHSQKVDQSOTIPKHVFE	960
PS	Claim 20; Page 3761-3764; 6221pp; English.	Qy	961	HTFHLRLEAGEYOQIMTASVSGSLKNQINNVGRTVPASVQGVVIADNAYSSYSLIVSWQKAA	1020
XX					
XX	The invention relates to polynucleotides (AAK51456-AAK53435) and the	Db	961	HTFHLRLEAGEYOQIMTASVSGSLKNQINNVGRTVPASVQGVVIADNAYSSYSLIVSWQKAA	1020
CC	encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to	Qy	1021	GVAERYDILLTENGILLRNTSEBATTQKHKFEDLTGKKYKIQILTVSGGLFSKEAQTE	1080
CC	cytokine, cell proliferation or cell differentiation or which may induce	Db	1021	GVAERYDILLTENGILLRNTSEBATTQKHKFEDLTGKKYKIQILTVSGGLFSKEAQTE	1080
CC	production of other cytokines in cell populations. The	Qy	1081	GRTVPAAVTDLRITENSTRHLSFRWTASEGELSWYNIFLYNPDGNLQERAQVDPLVQSFS	1140
CC	polynucleotides and polypeptides are useful in gene therapy, vaccines or	Db	1081	GRTVPAAVTDLRITENSTRHLSFRWTASEGELSWYNIFLYNPDGNLQERAQVDPLVQSFS	1140
CC	peptide therapy. The polypeptides have various cytokine-like activities,	Qy	1141	FQNLQGRMYKMWIVTHSGELSNESFIFGRTVPASVSHLRGSRNRTTDSLWFNWSPASGD	1200
CC	e.g. stem cell growth factor activity, haematopoiesis regulating	Db	1141	FQNLQGRMYKMWIVTHSGELSNESFIFGRTVPASVSHLRGSRNRTTDSLWFNWSPASGD	1200
CC	activity, tissue growth factor activity, immunomodulatory activity and	Qy	1201	PDFYELLILYNPNTGKKNWKDKOLTEWRFOGLVPGRYKVLVWVTHSGDLNKNVTAESRTA	1260
CC	activin/inhibin activity and may be useful in the diagnosis and/or	Db	1201	PDFYELLILYNPNTGKKNWKDKOLTEWRFOGLVPGRYKVLVWVTHSGDLNKNVTAESRTA	1260
CC	treatment of cancer, leukaemia, nervous system disorders, arthritis and	Qy	1261	PSPPSLMSFADIANSTSLAITWKGPDPDWDYNDPELOMLPRDALTVFNPYNNRKSSEGRIVY	1320
CC	inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111	Db	1261	PSPPSLMSFADIANSTSLAITWKGPDPDWDYNDPELOMLPRDALTVFNPYNNRKSSEGRIVY	1320
CC	(AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the	Qy	1321	GLRPGRSYQFNVTKSGDSWKTSKPIFGSVRTKPKDKIQNLHCRPQNSTAIACSWIPDPS	1380
CC	sequence listing were missing at the time of publication	Db	1321	GLRPGRSYQFNVTKSGDSWKTSKPIFGSVRTKPKDKIQNLHCRPQNSTAIACSWIPDPS	1380
XX					
XX	Sequence 1997 AA;	Qy	1381	DFDGYSECRMDTQEVFEFSRKEKEKSLINIMMLVPHKRYLVSIKVSQAGMTSEVVEDS	1440
XX					
Query Match	99.8%; Score 10457; DB 4; Length 1997;				
Best Local Similarity	99.8%; Pred. No. 0;				
Matches 1993; Conservative	1; Mismatches 3; Indels 0; Gaps 0;				
Qy	1	MLSHGAGLAWITLSLLOGLAEPCNFTLAESKASHSVSIQWRILGSCNFSLIYSS	60		
Db	1	MLSHGAGLAWITLSLLOGLAEPCNFTLAESKASHSVSIQWRILGSCNFSLIYSS	60		
Qy	61	DTLGAALCPTFRIDNTTYGCMQLDQAGTIYNFKIISLDEERTVVLQDPLPPARFGVSK	120		
Db	61	DTLGAALCPTFRIDNTTYGCMQLDQAGTIYNFKIISLDEERTVVLQDPLPPARFGVSK	120		
Qy	121	EKTTSTGLHVVMTTSSGKVTSEYQLFDENNQIKQGVQIOBSTSWNEYTFPNTAGSKYN	180		
Db	121	EKTTSTGLHVVMTTSSGKVTSEYQLFDENNQIKQGVQIOBSTSWNEYTFPNTAGSKYN	180		
Qy	181	IAITAVSGGKRSFVYTTNGSTVSPVKDIGISTKANSLIISWSHSGGNVRYRLMLMDKG	240		
Db	181	IAITAVSGGKRSFVYTTNGSTVSPVKDIGISTKANSLIISWSHSGGNVRYRLMLMDKG	240		
Qy	241	ILVHGGVVDKATSVAFHGLSPGYLYNLTVNTEAAGLQNYRWKLVRTAPMEVSNLKVTDN	300		
Db	241	ILVHGGVVDKATSVAFHGLSPGYLYNLTVNTEAAGLQNYRWKLVRTAPMEVSNLKVTDN	300		

Db 1381 DFDGYSIECRXMDTQVEFSRLEKEKSLNIMLVPHKRYLVSIKVQSGMTSEVVEDS 1440  
Qy 1441 TITMIDRPPPPHRIRVNEKDVLSKSSINFVNCWFSFDTNGAVKYFTVVVRADGSD 1500  
Db 1441 TITMIDRPPPPHRIRVNEKDVLSKSSINFVNCWFSFDTNGAVKYFTVVVRADGSD 1500  
Qy 1501 LKPEQOHPPLSYLEYRHNASIRVYQTNFYASKCAENPNSNSKSFNFKLGAEWESLGGKRD 1560  
Db 1501 LKPEQOHPPLSYLEYRHNASIRVYQTNFYASKCAENPNSNSKSFNFKLGAEWESLGGKRD 1560  
Qy 1561 PTQOKFCGPKPHATYRISRAFTQLPDEDLKFTTPLYSDFPSPITTESPLFGAI 1620  
Db 1561 PTQOKFCGPKPHATYRISRAFTQLPDEDLKFTTPLYSDFPSPITTESPLFGAI 1620  
Qy 1621 EGVSAAGFLIGMLVAVVALLICROKVSIGRPERPARLSIRDRPLSVHLNLGQGNKRTS 1680  
Db 1621 EGVSAAGFLIGMLVAVVALLICROKVSIGRPERPARLSIRDRPLSVHLNLGQGNKRTS 1680  
Qy 1681 CPIKINQFEGHFMKLQADSNVLLSKEYEELKDVGNQSCDIALLPENRGNRYNNILPYD 1740  
Db 1681 CPIKINQFEGHFMKLQADSNVLLSKEYEELKDVGNQSCDIALLPENRGNRYNNILPYD 1740  
Qy 1741 ATRVKLSNVDDPCSDYNASYPGNFRREYIVTQGPLPOTKDDFWKMWQVQNVHIVM 1800  
Db 1741 ATRVKLSNVDDPCSDYNASYPGNFRREYIVTQGPLPOTKDDFWKMWQVQNVHIVM 1800  
Qy 1801 VTQCVKGRVKCDHYWADQDSLYGDLILQMLSESVLPEWTFEKFICGEEQDAHRLI 1860  
Db 1801 VTQCVKGRVKCDHYWADQDSLYGDLILQMLSESVLPEWTFEKFICGEEQDAHRLI 1860  
Qy 1861 RHFTYTWPDHGVPETTQSLIQFVRTVDRDYNRSPGAGPTVHCSAGVGRGTGTFIALDRI 1920  
Db 1861 RHFTYTWPDHGVPETTQSLIQFVRTVDRDYNRSPGAGPTVHCSAGVGRGTGTFIALDRI 1920  
Qy 1921 LQQLDSKDSVDIYGAVHDLRLHRVHMVQTECOQYVYLHQCVRDVLRLARKLSEQNPLFPI 1980  
Db 1921 LQQLDSKDSVDIYGAVHDLRLHRVHMVQTECOQYVYLHQCVRDVLRLARKLSEQNPLFPI 1980  
Qy 1981 YENVNPEVHRDPVYSRH 1997  
Db 1981 YENVNPEVHRDPVYSRH 1997  
RESULT 7  
AAM79805  
ID AAM79805 standard; protein; 2002 AA.  
XX AC AAM79805;  
XX DT 06-NOV-2001 (first entry)  
XX DE Human protein SEQ ID NO 3451.  
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation.  
XX OS Homo sapiens.  
XX PN WO200157190-A2.  
XX PD 09-AUG-2001.  
XX PF 05-FEB-2001; 2001WO-US004098.  
XX PR 03-FEB-2000; 2000US-00496914.  
PR 27-APR-2000; 2000US-00560875.  
PR 20-JUN-2000; 2000US-00598075.  
PR 19-JUL-2000; 2000US-00620325.  
PR 01-SEP-2000; 2000US-00654936.  
PR 15-SEP-2000; 2000US-00663561.

PR 20-OCT-2000; 2000US-00693325.  
PR 30-NOV-2000; 2000US-00728422.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;  
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
XX WPI: 2001-476283/51.  
DR N-PSDB; AAK52938.  
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful  
PT in diagnosis and gene therapy.  
XX Claim 20; Page 340-341; 6221pp; English.  
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the  
CC sequence listing were missing at the time of publication  
XX SQ Sequence 2002 AA;

Query Match 99.6%; Score 10440.5; DB 4; Length 2002;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1992; Conservative 1; Mismatches 4; Indels 1; Gaps 1;  
Qy 1 MLSHGAGLAWITLSLLOTLGLAEPFCNFTLAESKASSHSVSQWRLISGSPCNFSLIYSS 60  
Db 5 MLSHGAGLAWITLSLLOTLGLAEPFCNFTLAESKASSHSVSQWRLISGSPCNFSLIYSS 64  
Qy 61 DTLGAALCPTFRIDNTTYGNCNLQDLQAGTIYNFKILSLDBERTVVLQTDPLPPARFVSK 120  
Db 65 DTLGAALCPTFRIDNTTYGNCNLQDLQAGTIYNFKILSLDBERTVVLQTDPLPPARFVSK 124  
Qy 121 EKTSTGLHVMWTPSSGKVTSEVOLFDENNQIKQVQIOESTSWNEYTFNLTAGSKYN 180  
Db 125 EKTSTGLHVMWTPSSGKVTSEVOLFDENNQIKQVQIOESTSWNEYTFNLTAGSKYN 184  
Qy 181 IAITAVSGGKRSPSVYTNSTGTPSPVKDIGISTKANSLISWGHSGNVERYRLMLMDKG 240  
Db 185 IAITAVSGGKRSPSVYTNSTGTPSPVKDIGISTKANSLISWGHSGNVERYRLMLMDKG 244  
Qy 241 ILVHGGVVDKHAITSYAFHGLSPGYLYNLVTMTAAGLQNYRWKLVRTAPMEVSNLKVTDND 300  
Db 245 ILVHGGVVDKHAITSYAFHGLSPGYLYNLVTMTAAGLQNYRWKLVRTAPMEVSNLKVTDND 304  
Qy 301 GSITSLSKVKWQRPNGVDSYNIITLSHGKTIKESRVLAPWITETHEKELVPGRLYQVTVSC 360  
Db 305 GSITSLSKVKWQRPNGVDSYNIITLSHGKTIKESRVLAPWITETHEKELVPGRLYQVTVSC 364  
Qy 361 VS-GELSAQKMAVGRTPFDKVAANLEANNNGRMESLVVSNPSPAGDWEQYRILLFNDSVVL 419  
Db 365 VSLGELSAQKMAVGRTPFDKVAANLEANNNGRMESLVVSNPSPAGDWEQYRILLFNDSVVL 424  
Qy 420 LNIITVGKEETQYVMDDTGLVPGRQYEVETVIVESGNLKNSERCQGRTPVLAQLRVKXAN 479  
Db 425 LNIITVGKEETQYVMDDTGLVPGRQYEVETVIVESGNLKNSERCQGRTPVLAQLRVKXAN 484  
Qy 480 ETSLSIMWQTPVAWEKYYIISLADRLILLHKSLSKDAKEFTTDLVPGHKYMATVTSIS 539  
Db 485 ETSLSIMWQTPVAWEKYYIISLADRLILLHKSLSKDAKEFTTDLVPGHKYMATVTSIS 544

QY 540 GDLKNSSVKGRTPVPAQVTDLHVANQGWTSLSLTNTWTOAQGDVEFYOVLLIHENNVIKNE 599  
DB 545 GDLKNSSVKGRTPVPAQVTDLHVANQGWTSLSLTNTWTOAQGDVEFYOVLLIHENNVIKNE 604  
QY 600 SISSETSRYSFHSLSKSLYSVVVTVTTGGISSRQVVVEGRTVPSSVGVTVNNSGRNDY 659  
DB 605 SISSETSRYSFHSLSKSLYSVVVTVTTGGISSRQVVVEGRTVPSSVGVTVNNSGRNDY 664  
QY 660 LSVSWLAPGDVDNVEYVTLSDHGKVVQSLVITAKSVRECSFSSLTTPGRLYTVTTITRSKY 719  
DB 665 LSVSWLAPGDVDNVEYVTLSDHGKVVQSLVITAKSVRECSFSSLTTPGRLYTVTTITRSKY 724  
QY 720 ENHSFSQERTVPDKVQVSVNSARSYDLRVSWHATGDPHYEVTIKKNNFIQTKSIP 779  
DB 725 ENHSFSQERTVPDKVQVSVNSARSYDLRVSWHATGDPHYEVTIKKNNFIQTKSIP 784  
QY 780 KSENECVFVQLVPGRLYSVTVTTKSGQYEAQNGRRTIPEPVKDLTLNRSTEDLHVTW 839  
DB 785 KSENECVFVQLVPGRLYSVTVTTKSGQYEAQNGRRTIPEPVKDLTLNRSTEDLHVTW 844  
QY 840 SGANGVDVQYIEIQLLFNDMKVPPFHLVNTATEYRFTSLTPGROYKILVLTISGDVQOSA 899  
DB 845 SGANGVDVQYIEIQLLFNDMKVPPFHLVNTATEYRFTSLTPGROYKILVLTISGDVQOSA 904  
QY 900 FIEGFTVPSAVKNIHISPGATDSLTVNWTGGGDVDSYTVSAFRHSQKVDSDQTIKPHVF 959  
DB 905 FIEGFTVPSAVKNIHISPGATDSLTVNWTGGGDVDSYTVSAFRHSQKVDSDQTIKPHVF 964  
QY 960 EHTFHRLEAGEQYQIMIASVSGSLKNOINNVGRTVPASVQGVIAADNAYSYSILTVSWOKA 1019  
DB 965 EHTFHRLEAGEQYQIMIASVSGSLKNOINNVGRTVPASVQGVIAADNAYSYSILTVSWOKA 1024  
QY 1020 AGVAERYDILLTENGILLRNSTSPATTQKHFPDLTPGKKYKIQIILTVSGGLFSPKSAQT 1079  
DB 1025 AGVAERYDILLTENGILLRNSTSPATTQKHFPDLTPGKKYKIQIILTVSGGLFSPKSAQT 1084  
QY 1080 EGRTPVPAVVDLRTENSTHLSFRWTASGELSWYNIFLVNPNGNLQERAQVDPLVQSF 1139  
DB 1085 EGRTPVPAVVDLRTENSTHLSFRWTASGELSWYNIFLVNPNGNLQERAQVDPLVQSF 1144  
QY 1140 SFQNLQGRYKMWIVTHSGELSNESIFGRTPVSPASVSHLGSNRNTTDSLWFNWPASG 1199  
DB 1145 SFQNLQGRYKMWIVTHSGELSNESIFGRTPVSPASVSHLGSNRNTTDSLWFNWPASG 1204  
QY 1200 DFDYFYLILYNPNTGKKNKODLTWRFGQLVPGRYKYLWVWTHSGDLSNKVTASRT 1259  
DB 1205 DFDYFYLILYNPNTGKKNKODLTWRFGQLVPGRYKYLWVWTHSGDLSNKVTASRT 1264  
QY 1260 APSPPSLMSFADIANSTSLAITWKGPPDWTYNDYFELQWLPDALTVPNPNYNNRSEGRIV 1319  
DB 1265 APSPPSLMSFADIANSTSLAITWKGPPDWTYNDYFELQWLPDALTVPNPNYNNRSEGRIV 1324  
QY 1320 YGLRPGRSYQFNKVTSGDSWKTSKPIFGSVRTKPKDKIQLNHLCPQNSTAIACSWIPPD 1379  
DB 1325 YGLRPGRSYQFNKVTSGDSWKTSKPIFGSVRTKPKDKIQLNHLCPQNSTAIACSWIPPD 1384  
QY 1380 SDFDGYIECRKMDTQVEFSRKLKESKSLINIMLVPHKRYLSIKVQSGAGMTSEYVED 1439  
DB 1385 SDFDGYIECRKMDTQVEFSRKLKESKSLINIMLVPHKRYLSIKVQSGAGMTSEYVED 1444  
QY 1440 STITMIDRPPPPPHIRHNEKDVLSKSSINFTVNCWFSDTNGAVKYFTVVVREADGSD 1499  
DB 1445 STITMIDRPPPPPHIRHNEKDVLSKSSINFTVNCWFSDTNGAVKYFTVVVREADGSD 1504  
QY 1500 ELKEPQOHPPLSYLEYRHNASIRVQYNYFASKCAENPNNSKSFNFKLGAEMSLGGR 1559  
DB 1505 ELKEPQOHPPLSYLEYRHNASIRVQYNYFASKCAENPNNSKSFNFKLGAEMSLGGR 1564  
QY 1560 DPTQOQKCDGPKLPHYATYRISIRAFQTLFDBDLKEFTKPLYSDDTFPSLPTTSEPLFGA 1619  
DB 1565 DPTQOQKCDGPKLPHYATYRISIRAFQTLFDBDLKEFTKPLYSDDTFPSLPTTSEPLFGA 1624  
QY 1620 IEGVSAGLFLIGMLVAVVALLICRQKVSHGRERPSARLSIRDRPLSVHLNLQKGNRKT 1679

DB 1625 IEGVSAGLFLIGMLVAVVALLICRQKVSHGRERPSARLSIRDRPLSVHLNLQKGNRKT 1684  
QY 1680 SCPIKINQFEGHFMKLOADSNYLLSKEYBELKDVGRNQSCDIALLPENRGNRYNNILPY 1739  
DB 1685 SCPIKINQFEGHFMKLOADSNYLLSKEYBELKDVGRNQSCDIALLPENRGNRYNNILPY 1744  
QY 1740 DATRVKLSNVDDPCSDYINASYIPGNPFREYITVQGPLGTGKDDFMKMWQNVNINIV 1799  
DB 1745 DATRVKLSNVDDPCSDYINASYIPGNPFREYITVQGPLGTGKDDFMKMWQNVNINIV 1804  
QY 1800 MVTQCVKEGRVKCDHWPADQDSLYYGLLILQMLSESVLPWTIREFKICGEEQDLAHL 1859  
DB 1805 MVTQCVKEGRVKCDHWPADQDSLYYGLLILQMLSESVLPWTIREFKICGEEQDLAHL 1864  
QY 1860 IRHFHYTVPDHPGVPETTSQSLIOFVRTVDYINRSPGAGTVPVHCSAGVGRGTGFIALDR 1919  
DB 1865 IRHFHYTVPDHPGVPETTSQSLIOFVRTVDYINRSPGAGTVPVHCSAGVGRGTGFIALDR 1924  
QY 1920 ILQOLDSKDSVDIYGAVHDLRLHRVHMVQTECOYVYLHQCVRDVLARKLRSEQENPLFP 1979  
DB 1925 ILQOLDSKDSVDIYGAVHDLRLHRVHMVQTECOYVYLHQCVRDVLARKLRSEQENPLFP 1984  
QY 1980 IYENVNPEYHRDPVYSRH 1997  
DB 1985 IYENVNPEYHRDPVYSRH 2002

RESULT 8  
AAO24268  
ID AAO24268 standard; protein; 1998 AA.  
XX AC AAO24268;  
XX DT 11-MAR-2004 (first entry)  
XX DE Murine vascular endothelial-protein tyrosine phosphatase.  
XX KW Vascular endothelial protein-tyrosine phosphatase; VE-PTP; mouse; human;  
XX KW gene therapy; cytosolic; VE-cadherin; vascular endothelial-cadherin.  
XX OS Mus sp.  
XX PN WC2003084565-A2.  
XX PD 16-OCT-2003.  
XX PF 08-APR-2003; 2003WO-EP003645.  
XX PR 08-APR-2002; 2002EP-00007837.  
XX PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
XX PI Nawroth R, Deutsch U, Vestweber D, Shima DT, Golding M;  
XX DR WPI; 2003-804251/75.  
XX PT Use of the polypeptide comprising vascular endothelial-protein tyrosine  
XX PT phosphatase (VE-PTP) or the nucleic acid encoding the polypeptide for the  
XX PT manufacture of an agent for monitoring or modulating VE-cadherin mediated  
XX PT disorders.  
XX PS Example; Fig 2; Opp; English.  
XX CC The present invention relates to a polypeptide comprising vascular  
XX CC endothelial-protein tyrosine phosphatase (VE-PTP) or its active fragment  
XX CC or effector, or the nucleic acid encoding the polypeptide or its  
XX CC effector, for use in the manufacture of an agent for monitoring or  
XX CC modulating VE-cadherin mediated processes or disorders. The polypeptide  
XX CC comprising vascular endothelial-protein tyrosine phosphatase (VE-PTP) or  
XX CC its active fragment or effector, or the nucleic acid encoding the  
XX CC polypeptide or its effector, is useful for the manufacture of an agent  
XX CC for monitoring or modulating VE-cadherin mediated processes or disorders.



CC e.g., cancer. The present sequence is a protein shown in the  
CC exemplification of the invention  
XX  
SQ Sequence 1998 AA;

Query Match 87.0%; Score 9117.5; DB 7; Length 1998;  
Best Local Similarity 85.7%; Pred. No. 0;  
Matches 1713; Conservative 129; Mismatches 155; Indels 1; Gaps 1;

Qy	1	MLSHGAGLAWITLSLQTLGAEPKCNFTLAESKASHSVSIOWRILGSPCNFSLIYSS	60
Db	1	MLRHGALTALWITUSVQVQAEQKCNFTLLESRVSSLSASIQWRTFASPCNFSIYSS	60
Qy	61	DTLGAALCPTFRIDNTYTCNQLQDLQAGTIYNFKIISLD-EERTVWLQTDPLPAPRGVS	119
Db	61	DTSGPMCHPIRIDNTFYGCNPKDLQAGTVYNFIVSLDGEESDVLQTDPLPAPFEVN	120
Qy	120	KEKTSITGLHVMWTPSSGKVTSYEVQLFDENNQKIQGVQIQESTWMEYTFNLTAGSKY	179
Db	121	PEKTASTTLQVRWTPSSGKVSWEVQLFDHNNQKIQEVQVQESTWTSQYTFNLTEGNSY	180
Qy	180	NIAITAVSGGKRSFVNTGSTVSPVKDIGISTKANSLISWSHSGSNVERELMLMDK	239
Db	181	KVAITAVSGEKRSPVYINGSTVSPVKDIGISPNPNSLLISWSRSGSNVEQVRLVMDK	240
Qy	240	GILVHGVDVKHATSYAFHGLSPGYLYNLVTMTAAAGLQNYRWKLVRTAPMEVSNLKVTN	299
Db	241	GAIQDITNVDKRTSYAFHELTTPGHLNLTIVTWASGLQNSRWKLVRTAPMEVSNLKVTN	300
Qy	300	DGSLTSLVKVQRPPGNVDSNITLSHGKTIKESRVLAPWITETHPKELVPGRLYQVTS	359
Db	301	DGRLTSLNVKQKPPGVDVSYITLSHGKTIKESKTLAPVTEQFKDLVPGRLYQVTS	360
Qy	360	CVSGELSAQKVAWGRTTPDKVANLEANNGMRSLVSVSPAGDWEQYRILLPNDSVVL	419
Db	361	CISGELSAKSAAGRTVPEKVRNLVSYNEIWMKSFTVNTTPPAGDWEHYRILVFNESLVL	420
Qy	420	LNITVGKEETQYVMDTGLVPGROYEVEIVESGNLKNRSCQRTVPLAVLQVRKHAN	479
Db	421	LNITVGKEETHYALDGLLEIPGRQYEIEVIVESGNLKNRSCQRTVPLAVLQVRKHAN	480
Qy	480	ETSLSIMQTPVAWEKYYIISLADRDLLIHKSLSKDAKEFTFTDLVPGRYKMAVTSIS	539
Db	481	ETSLGITWRAPLGEWEKYYIISLMDRELLVIHKSLSKDAKEFTFTDLMPGRNYKATVTSMS	540
Qy	540	GDLSKSSSVKGRTPVAQVTDLHVANQGTSSLFNWTQAQGDVEFYQVLLIHENVVYKNE	599
Db	541	GDLSKSSSVKGRTPVAQVTDLHVANQGTSSLFNWTQAQGDVEFYQVLLIHENVVYKNE	600
Qy	600	SISSETSRYSFHLKSGSLSVVTVTSGGSISSRQVVEGRTVPSSVSGVTVVNSGRNDY	659
Db	601	SVSSDTSRYGFRAKLPGLSVVTVTSGGSISSRQVVAEGRTVPSSVSGVTVVNSGRNDY	660
Qy	660	LSVSWLVPAGDVNDYVTLSDHGKVSQSLVIKSVRECSFSLTPGRLYTVTITRSKY	719
Db	661	LSVSWLVPAGDVNDYVTLSDHGKVSQSLVIKSVRECSFSLTPGRLYTVTITRSKY	720
Qy	720	ENHSFSGERTVPKQVQSVNSNARSYLAVSWHATGDFDHYEVTIKKNQNTIQKSIIP	779
Db	721	ASHSFTBERTVPKQVQSVNSNARSYLAVSWHATGDFDHYEVTIKKNQNTIQKSIIP	780
Qy	780	KSENECVFVLQVRLYSVTVTKSGQYEAQNGQRTIPEPVKDLTLNRKSTEDLHVTW	839
Db	781	KSENECEFIELVPGRLYSVTVTKSGQYEAQNGQRTIPEPVKDLTLNRKSTEDLHVTW	840
Qy	840	SGANGVDVQYEIQLLFNDMKVPPPPHVLNVTATEYRFTSLTPGRQYKILVLTISGDVQOSA	899
Db	841	SRANGVDVQYEVQLLFNDMKVFPFPHILVNTATEYKFTALTTPGRHYKILVLTISGDVQOSA	900
Qy	900	FIEGFTVPSAVKNIHISPNKATDSLNTVNTWTPGGDVDSYTVSAFRHSQKVDSTQIPKHFV	959
Db	901	FIEGLFVPSTVTKNIHISANGATDRLMVTWSPGGDVDSYTVSAFRHSQKVDSTQIPKHFV	960

Qy	960	EHTFHRLEAGEQOQIMIASVSGSLKQINQVWTRVTPASVQGVIAADNAYSYSLSIVSWQKA	1019
Db	961	EHTFHRLEAGAKYRIAIVSVSGSLRNOIDALQOTVTPASVQGVVAANAYSNSLSIVSWQKA	1020
Qy	1020	AGVAERYDIIILLTENGILLRNTSEPAATKOHKFDLTPGKKYKIQLITVSGGLFSKBAQT	1079
Db	1021	LGAERYDIIILLNENGILLRNTSEPAATKOHKFDLTPGKKYKQILITVSGGLFSKESQA	1080
Qy	1080	EGRTVPAAVTDLRIITENSTRHLSFRWTASEGELSWYNIFLYNPDGNIQERAOVDPLVQSF	1139
Db	1081	EGRTVPAAVTNLRIITENSTRHLSFRWTASEGELSWYNIFLYNPDRTIQERAOVDPLVQSF	1140
Qy	1140	SFQNLQGRMYKMWIVTHSGELSNESIFIGRTVPASVSHLRGSRNRTTDSLWFNWSPASG	1199
Db	1141	SFQNLQGRMYKMWIVTHSGELSNESIFIGRTVPAAVNHKLGSHRNTTDSLWFNWSPASG	1200
Qy	1200	DFDFYELILYNPNTGKKNWKDOLTEWRFOGLVPGRYKYLWVYVTHSGDLSNKTABSR	1259
Db	1201	DFDFYELILYNPNTGKKNWKEDVTEWRFOGLVPGRYKYLWVYVTHSGDLSNKTABSR	1260
Qy	1260	APSPSLMSFADTANTSLAITWKGPDPDWDYNDPELOQLPRDALTVENPNYNNRSEGRIV	1319
Db	1261	APSPSLMSFADVANTSLAITWKGPDPDWDYNDPELOQFPGDALTIENPNYNNRSEGRIV	1320
Qy	1320	YGLRPGRSYQFNKTVSGDSWKTYSKPIFGSVRTKPKIQNLHCRPQNSTAIACSWIPPD	1379
Db	1321	YGLRPGRSYQFVKTVSGDSWKTYSKPIFGSVRTKPKIQNLHCRPQNSTAIACSWIPPD	1380
Qy	1380	SDFGYSIECRKMDTOEVEFSRKLKESKLSLNTIMMLVPHKRYLVSIVKQAGMTSEVED	1439
Db	1381	SDFGYSIECRKMDTOEIEFSRKLKESKLSLNTIMMLVPHKRYLVSIVKQAGMTSEVED	1440
Qy	1440	STITMDIRPPPPPHIRVNEKDVLISSKSNFTVNCNWSFSDTGAVKYFTVVVREADGSD	1499
Db	1441	STITMDIRPPPPPHIRVNEKDVLISSKSNFTVNCNWSFSDTGAVKYFAVVVREADGSD	1500
Qy	1500	ELKPEQOHPPLPSYLEYRHNASIRVQNTYPASKCAENPNNSKSNFNIKLAEMESLGKGR	1559
Db	1501	ELKPEQOHPPLPSYLEYRHNASIRVQNTYPASKCAENPNNSKSNFNIKLAEMESLGKGR	1560
Qy	1560	DPTQKFCDCGPLKPHYATYRISIRAFQTFDEDLKFTKPLYSIDTFPSLPTTSESEPLFGA	1619
Db	1561	DPSQKFCDCGPLKPHYATYRISIRAFQTFDEDLKFTKPLYSIDTFPSLPTTSESEPLFGV	1620
Qy	1620	IEGVASGLFLIGMLVAVALLICROKVSCHGRERPSARLSIRDRPLSVHLNLGOKGNRKT	1679
Db	1621	IEGVASGLFLIGMLVALVAFFICROKASHRERPSARLSIRDRPLSVHLNLGOKGNRKT	1680
Qy	1680	SCPIKINQFEGHPMKLQADSNYLLSKEYEELKDVGRNQSCDIALPENRGNKNYNNILPY	1739
Db	1681	SCPIKINQFEGHPMKLQADSNYLLSKEYEELKDVGRNQSCDIALPENRGNKNYNNILPY	1740
Qy	1740	DATRVKLSNVDDPCSDYINASYIPGNPFREYIVTQGPLPGTKDDPFKMWQVNSQVHNIV	1799
Db	1741	DASRVKLSNVDDPCSDYINASYIPGNPFREYIVTQGPLPGTKDDPFKMWQVNSQVHNIV	1800
Qy	1800	MVTQCEKGRVKCDHWAPDODSLYYGDLTLOMLSESVLPEWTTIRBFKICSEBOLDARHL	1859
Db	1801	MVTQCEKGRVKCDHWAPDQDPLYYGDLTLOMLSESVLPEWTTIRBFKICSEBOLDARHL	1860
Qy	1860	IRHFHYTVWPDHGVPTTQSLIQFVRTVDYINRSPGAGTVVHCAGVGTGTFTALDR	1919
Db	1861	IRHFHYTVWPDHGVPTTQSLIQFVRTVDYINRSPGAGTVVHCAGVGTGTFTALDR	1920
Qy	1920	ILQOLSKOSVDIYGAVHDLRLHRVHMVQTECOYVYLHQCVRDVLRLARKLRSQENPLFP	1979
Db	1921	ILQOLSKOSVDIYGAVHDLRLHRVHMVQTECOYVYLHQCVRDVLRLARKLRSQENPLFP	1980
Qy	1980	IYENNVPEYHRDPVYSRH 1997	
Db	1981	IYENNVPEYHRDAIYSRH 1998	





Db 961 SDTNGAVKFTVVVREADGSDLPQHQHPSPSYLEVRHNASIRVYQNTYFASKCAENPN 1020  
Qy 1539 SNSKSFNIKGAEMESLGGKEDPTQOKFCDCGPKPHTAYRISIRAFQOLFDEDLKEFTKP 1598  
Db 1021 SNSKSFNIKGAEMESLGGKCDPTQOKFCDCGPKPHTAYRISIRAFQOLFDEDLKEFTKP 1080  
Qy 1599 LYSTFTFSLPTTSESEPLFGAIEGVSAGLFLIGMLVAVVALLICQKVSHGRERPSARLS 1658  
Db 1081 LYSTFTFSLPTTSESEPLFGAIEGVSAGLFLIGMLVAVVALLICQKVSHGRERPSARLS 1140  
Qy 1659 IRRDRPLSVHLNLGCKGNKRTSCPIKNOPEGHFMKLQADSNYLLSKEYEELKDVGRNOS 1718  
Db 1141 IRRDRPLSVHLNLGCKGNKRTSCPIKNOPEGHFMKLQADSNYLLSKEYEELKDVGRNOS 1200  
Qy 1719 CDIALLENKRNKYNILPYDATRVKLSNVDDPCSDYINASVPGNFRREYIVTQGP 1778  
Db 1201 CDIALLENKRNKYNILPYDATRVKLSNVDDPCSDYINASVPGNFRREYIVTQGP 1260  
Qy 1779 LPGTKDFWKMVQBNVHNIWMVTCQVEKGRVKCDHYWPADQDSLYYGDILQMLSESVL 1838  
Db 1261 LPGTKDFWKMVQBNVHNIWMVTCQVEKGRVKCDHYWPADQDSLYYGDILQMLSESVL 1320  
Qy 1839 PEWTIREFKICGEQOLDARHLIRHFHYTVMPDHGVPETTQSLIOFVRTVDRDYNRSPGAG 1898  
Db 1321 PEWTIREFKICGEQOLDARHLIRHFHYTVMPDHGVPETTQSLIOFVRTVDRDYNRSPGAG 1380  
Qy 1899 PTVHCSAGVGRGTFTFIALDRILQOLDSKDSVDIYGAVHDLRLHRVHMVOTECQYVYLHQ 1958  
Db 1381 PTVHCSAGVGRGTFTFIALDRILQOLDSKDSVDIYGAVHDLRLHRVHMVOTECQYVYLHQ 1440  
Qy 1959 CVRDVLRARK 1968  
Db 1441 CVRDVLRARK 1450

RESULT 10  
ADR10246

ID ADR10246 standard; protein; 1407 AA.

AC ADR10246;

DT 04-NOV-2004 (first entry)

DE Human protein useful for treating neurological disease Seq 3752.

KW human; oligo-capping method; diagnostic marker; gene therapy;

KW osteoporosis; neurological disease; Alzheimer's disease;

KW Parkinson's disease; dementia; short memory; cancer;

KW sense or motor function; emotional reaction; fear response; panic;

KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytoskeletal;

OS Homo sapiens.

PN EP1447413-A2.

PP 16-AUG-2004.

XX 12-FEB-2004; 2004EP-00003145.

XX 14-FEB-2003; 2003JP-00102207.

XX 09-MAY-2003; 2003JP-00131452.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;

PI Wakamatsu A, Ishii S, Nagai K, Irie R;

XX WPI; 2004-583265/57.

XX N-PSDB; ADR08290.

PT New 1995 cDNA, useful for treating osteoporosis, neurological diseases,

PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.  
XX Claim 1; SEQ ID NO 3752; 2686pp; English.

PS This invention relates to novel, isolated full length human cDNA  
XX molecules and the encoded proteins thereof. Specifically, it refers to  
CC cDNA clones obtained by an oligo-capping method, where none of these  
CC clones are identical to any known human mRNAs. The present invention  
CC describes an immunosay to identify agonists and antagonists, as well as  
CC antibodies, antisense molecules and siRNAs that can all be used to bind  
CC to and modulate expression of the cDNA molecules. As such, these  
CC molecules are useful for diagnostic markers or therapeutic targets for  
CC the various diseases or morbid states. In particular, they are useful in  
CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's  
CC disease, Parkinson's disease, dementia, short memory and various cancers,  
CC as well as for maintaining equilibrium of sense or motor function, and  
CC for treating emotional reaction, fear response and panic. Accordingly,  
CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,  
CC cytoskeletal and tranquiliser activities. This polypeptide is a protein  
CC encoded by a full length human cDNA sequence of the invention. NOTE: This  
CC sequence is not given in the sequence listing of the specification but  
CC can be obtained on CD-ROM from the European Patent Office, Vienna Sub-  
CC office.

XX SQ Sequence 1407 AA;

Query Match 64.0%; Score 6712; DB 8; Length 1407;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1288; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 18 QTGLAEPERCNFTLAESKASHSVSIQWRILGSPCNFSLIYSDTLGAALCPTFRIDNTT 77

Db 115 ETGLAEPERCNFTLAESKASHSVSIQWRILGSPCNFSLIYSDTLGAALCPTFRIDNTT 174

Qy 78 YGCNLODLQAGTIYNFKIISLDEERTVWLOTDPFARFVGSKETSTSLHVMWTPSSG 137

Db 175 YGCNLODLQAGTIYNFKIISLDEERTVWLOTDPFARFVGSKETSTSLHVMWTPSSG 234

Qy 138 KVTSEVOLDFENNOKIQGVQIOESTSWNEYTFNLTAGSKYNIATFVSGGKRSRVYT 197

Db 235 KVTSEVOLDFENNOKIQGVQIOESTSWNEYTFNLTAGSKYNIATFVSGGKRSRVYT 294

Qy 198 NGSTVSPVKDIGISTKANSLLISWSHGSGNVRIRMLMDKGLVHGGVVDKHTSYAF 257

Db 295 NGSTVSPVKDIGISTKANSLLISWSHGSGNVRIRMLMDKGLVHGGVVDKHTSYAF 354

Qy 258 HGLSPGYLYNLTVMTEAAGLQNYRWKLVRTAPMEVSNLKVNDGSLTSLKVKQRPFGNV 317

Db 355 HGLTPGYLYNLTVMTEAAGLQNYRWKLVRTAPMEVSNLKVNDGSLTSLKVKQRPFGNV 414

Qy 318 DSYNITLSHKGTIKESRVLAPWITETHFKELVPGRLVQVTVSCVSGELSAQKVAVGRTFP 377

Db 415 DSYNITLSHKGTIKESRVLAPWITETHFKELVPGRLVQVTVSCVSGELSAQKVAVGRTFP 474

Qy 378 DKVANLEANNGRMRSILVSWSPAGDWEQYRILLFNDSVYLLNITVGEETQYVMDDTG 437

Db 475 DKVANLEANNGRMRSILVSWSPAGDWEQYRILLFNDSVYLLNITVGEETQYVMDDTG 534

Qy 438 LVPGRQYEVETVIVESGNLKNSERCQGRTPVLAVQLRVKHANETSLSIMMOTPAVEWEKY 497

Db 535 LVPGRQYEVETVIVESGNLKNSERCQGRTPVLAVQLRVKHANETSLSIMMOTPAVEWEKY 594

Qy 498 IISLADRDLLLIHKSLSKDAKBEFTFDLPGRKYMATVTSISGDLKNSSSVKGRTPVAQV 557

Db 595 IISLADRDLLLIHKSLSKDAKBEFTFDLPGRKYMATVTSISGDLKNSSSVKGRTPVAQV 654

Qy 558 TDLHVNQGMWTSLSFTNTWTAQGDVEFYQVLLIHENVVKNESISSETSRYSHLSKSGS 617

Db 655 TDLHVNQGMWTSLSFTNTWTAQGDVEFYQVLLIHENVVKNESISSETSRYSHLSKSGS 714

Qy 618 LYSVWVTVTSSGGISSRQVVEGRTVPSVSGVTVNNSGRNDYLSVSWLVAPOVDNVEVT 677

Db 715 LYSVWVTVTSSGGISSRQVVEGRTVPSVSGVTVNNSGRNDYLSVSWLVAPOVDNVEVT 774

QY 678 LSHDGKQVQSLVIAKSVRECSFSLTPCRLTYVTITTRSGKYENHSQERTVPDKVQGV 737  
 Db 775 LSHDGKQVQSLVIAKSVRECSFSLTPCRLTYVTITTRSGKYENHSQERTVPDKVQGV 834  
 QY 738 SVNSARSADYLRVSWVHATGDFHVEVTIKNNFIOTKSPKSENECVFQVLPGRLYS 797  
 Db 835 SVNSARSADYLRVSWVHATGDFHVEVTIKNNFIOTKSPKSENECVFQVLPGRLYS 894  
 QY 798 VVTVTKSQYBANQNGRTTIPVVKOLTLNRSTEDLHVTSWANGDVQYEQIQLFND 857  
 Db 895 VVTVTKSQYBANQNGRTTIPVVKOLTLNRSTEDLHVTSWANGDVQYEQIQLFND 954  
 QY 858 MKVPFPFHLVNTATEYRFTSLTPGROYKILVLTISGDVQSAFTEGFTVPSAVKNIHSP 917  
 Db 955 MKVPFPFHLVNTATEYRFTSLTPGROYKILVLTISGDVQSAFTEGFTVPSAVKNIHSP 1014  
 QY 918 NGATDSLTVNTPGGDVDSVTVSAFRHSQKVDSTIPKHFVFEHFRLEAGEQYQIMIA 977  
 Db 1015 NGATDSLTVNTPGGDVDSVTVSAFRHSQKVDSTIPKHFVFEHFRLEAGEQYQIMIA 1074  
 QY 978 SVSGSLKNQINNVGRTVPASVQGVVIADNAYSYSILVSWQKAAGVABRYDILLTENGIL 1037  
 Db 1075 SVSGSLKNQINNVGRTVPASVQGVVIADNAYSYSILVSWQKAAGVABRYDILLTENGIL 1134  
 QY 1038 LRNTSEPAATTKQHKPDELTPGKKYKIQLITVSGGLFSKEAQTEGRTVPAAVTDLRITENS 1097  
 Db 1135 LRNTSEPAATTKQHKPDELTPGKKYKIQLITVSGGLFSKEAQTEGRTVPAAVTDLRITENS 1194  
 QY 1098 TRHLSFRWTA SEGELSWNIFLYNPDGNLOBRAQVDPLVQSFQNLQGRMYKQWIVTH 1157  
 Db 1195 TRHLSFRWTA SEGELSWNIFLYNPDGNLOBRAQVDPLVQSFQNLQGRMYKQWIVTH 1254  
 QY 1158 SGELSNESFIFGRTPVAPSVHLSRNSRNTTDSLFWNMPASGDFDFVELILYNPNGTKE 1217  
 Db 1255 SGELSNESFIFGRTPVAPSVHLSRNSRNTTDSLFWNMPASGDFDFVELILYNPNGTKE 1314  
 QY 1218 NWKDKLTWRPQGLVGRKYLWVTHSGDLSNKTVAESRTAPSPSLMSFADIANSTL 1277  
 Db 1315 NWKDKLTWRPQGLVGRKYLWVTHSGDLSNKTVAESRTAPSPSLMSFADIANSTL 1374  
 QY 1278 AITWKGPPDWDYNDYDFELOWLPDALTVPFNPN 1310  
 Db 1375 AITWKGPPDWDYNDYDFELOWLPDALTVPFNPN 1407  
 RESULT 11  
 ID AAB19773 standard; protein; 579 AA.  
 AC AAB19773;  
 XX 19-FEB-2001 (first entry)  
 DT 19-FEB-2001 (first entry)  
 XX Mouse vascular-endothelial protein tyrosine phosphatase.  
 DE Vascular-endothelial protein tyrosine phosphatase; VE-PTP; mouse; Tie-2;  
 KW receptor-type tyrosine kinase; antiangiogenic; antitumour;  
 KW antimetastatic; tumour; metastasis; angiogenesis; therapy.  
 XX Mus musculus.  
 OS Mus musculus.  
 XX Key Location/Qualifiers  
 FH Domain 3..24  
 FT /note= "membrane proximal fibronectin II domain"  
 FT 204..223  
 FT /note= "transmembrane domain"  
 FT 304..549  
 FT /note= "catalytic domain"  
 XX EP1046715-A1.  
 PN 25-OCT-2000.  
 XX

XX 23-APR-1999; 99EP-00108074.  
 XX 23-APR-1999; 99EP-00108074.  
 XX (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
 XX Fachinger G, Risau B, Deutsch U;  
 XX WPI; 2000-648932/63.  
 XX N-PSDB; AAA88865.  
 XX Monitoring or modulating Tie-2 tyrosine kinase activity, useful e.g. for  
 PT regulating tumor growth, using vascular-endothelial protein tyrosine  
 FT phosphatase.  
 XX Disclosure; Page 10-12; 60pp; English.  
 XX The present sequence is that of murine vascular-endothelial protein  
 CC tyrosine phosphatase (VE-PTP). VE-PTP is a member of subclass III  
 CC receptor type PTPs, bearing fibronectin type III-like repeats in the  
 CC extracellular domain and a single catalytic domain in the cytoplasmic  
 CC tail. VE-PTP specifically interacts with receptor-type tyrosine kinase  
 CC Tie-2, modulating its tyrosine phosphorylation. Tie-2 is involved in  
 CC angiogenic processes, the formation of blood vessels during embryonal  
 CC development, wound healing and in pathological processes such as tumour  
 CC development. VE-PTP or its catalytic domain, nucleic acids and ligands  
 CC can be used to monitor, stimulate or repress Tie-2 activity for the  
 CC purpose of monitoring or modulating angiogenesis, inducing or inhibiting  
 CC vascular growth or remodelling and blood vessel maturation, and  
 CC inhibiting tumour growth and metastasis  
 XX Sequence 579 AA;  
 SQ  
 Query Match 28.0%; Score 2937; DB 3; Length 579;  
 Best Local Similarity 94.3%; Pred. No. 4.3e-198;  
 Matches 546; Conservative 15; Mismatches 18; Indels 0; Gaps 0;  
 QY 1419 KRYLSIKVQSGAGTSEVEDSTITMIDRPPPPHPIRVNEKDVLSKSSINFTVNCSWF 1478  
 Db 1 KRYLSIKVQSGAGTSEVEDSTITMIDRPPPPHPIRVNEKDVLSKSSINFTVNCSWF 60  
 QY 1479 SDTNGAVKYTVVREADGSDDELKPEQOHPPLPSYLEYRHNASIRVYQTNYPASKCAENPN 1538  
 Db 61 SDTNGAVGYFAVVVREADSMDELKPEQOHPPLPSYLEYRHNASIRVYQTNYPASKCAENPN 120  
 QY 1539 SNSKSFNIKLGAEWESLGGKRDPTQOKFCGPKLPHPTAYRISIRAFQLPDDEDLKETKP 1598  
 Db 121 SSSKSFNIKLGAEWESLGGKRDPTQOKFCGPKLPHPTAYRISIRAFQLPDDEDLKETKP 180  
 QY 1599 LYSDFTFSPITTESEPLFGAIEGVSAGLFLIGMLVAVALLICRQKVSHGRERPSARLS 1658  
 Db 181 LYSDFTFSPITTESEPLFGVIEGVSAGLFLIGMLVAVALLICRQKVSHGRERPSARLS 240  
 QY 1659 IRDRPLSVHLNLGQGNRTKSCPIKINQEPHGMKLQADSNYLLSKSEYELKDVGRNQ 1718  
 Db 241 IRDRPLSVHLNLGQGNRTKSCPIKINQEPHGMKLQADSNYLLSKSEYELKDVGRSQ 300  
 QY 1719 CDITALLPENRGRYNNILPYDATRVKLSDVDDPCSDYINASYIPGNFRREYIVTQGP 1778  
 Db 301 CDITALLPENRGRYNNILPYDASRVKLSNVDDPCSDYINASYIPGNFRREYIVTQGP 360  
 QY 1779 LPGTKDDFWKMWVQNVHNIWMVTQCVKGRVKCDHYWAPADQDLSYTGDLILQMLSESVL 1838  
 Db 361 LPGTKDDFWKMWVQNVHNIWMVTQCVKGRVKCDHYWAPADQDLYYTGDLILQMLSESVL 420  
 QY 1839 PEWTIREFKICGEEQLDAHLIRHFHYTVMPDGVGPETTSQSLIQFVTRVDYINRSPGAG 1898  
 Db 421 PEWTIREFKICGEEQLDAHLIRHFHYTVMPDGVGPETTSQSLIQFVTRVDYINRSPGAG 480  
 QY 1899 PTVVHCSAGVGRGTGTFALDRILQQLDSDKSDYIYGAVHDLRLHRVHMVQTECOYVYLHQ 1958  
 Db 481 PTVVHCSAGVGRGTGTFALDRILQQLDSDKSDYIYGAVHDLRLHRVHMVQTECOYVYLHQ 540

```

Qy 1959 CVRDLRARKLRSEQENFLPFIYENVNPEYHRDPVYSRH 1997
Db 541 CVRDLRARKLRSEQENFLPFIYENVNPEYHRDAIYSRH 579

RESULT 12
ADQ66828
ID ADQ66828 standard; protein; 609 AA.
XX
AC ADQ66828;
XX
DT 07-OCT-2004 (first entry)
XX
DE Novel human protein sequence #1801.
XX
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.
XX
OS Homo sapiens.
XX
PN EPI440981-A2.
XX
PD 28-JUL-2004.
XX
PF 21-JAN-2004; 2004EP-00001196.
XX
PR 21-JAN-2003; 2003JP-00102206.
PR 09-MAY-2003; 2003JP-00131392.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Nagai K, Irie R;
XX
WPI; 2004-535376/52.
DR N-PSDB; ADQ64640.
XX
XX
Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS Claim 1; SEQ ID NO 3989; 2449pp; English.
XX
XX
The invention relates to 2495 novel polynucleotides (I) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to
CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC dementia and various cancers. This sequence corresponds to a protein
CC sequence of the invention.
XX
SQ Sequence 609 AA;

Query Match 17.7%; Score 1854.5; DB 8; Length 609;
Best Local Similarity 93.0%; Pred. No. 1.4e-121;
Matches 356; Conservative 5; Mismatches 3; Indels 19; Gaps 1;

Qy 18 QTGLAEPCNFTLAESKASHSVSIQWRIILGSPCNFSLIYSSDTLGAALCPTFRIDNTT 77
Db 236 ETGLAEPCNFTLAESKASHSVSIQWRIILGSPCNFSLIYSSDTLGAALCPTFRIDNTT 295
Qy 78 YGCNLQDLQAGTIYNFRIISLDEERTVVLQTDPLPPARFGVSKETKTTSTGLHVMWTPSSG 137
Db 296 YGCNLQDLQAGTIYNFRIISLDEERTVVLQTDPLPPARFGVSKETKTTSTGLHVMWTPSSG 355
Qy 138 KVTSEYVQLFDENNQKIQGVQIQBSTSWNEYTFPNLTAGSKYNIATVAVSGGRKRSFVYT 197
Db 356 KVTSEYVQLFDENNQKIQGVQIQBSTSWNEYTFPNLTAGSKYNIATVAVSGGRKRSFVYT 415

198 NGSTVPSPVKDIGISTKANSLIISWHSKSGNVERYLMLMDKGLVHGGVVDKHTSYAF 257
416 NGSTVPSPVKDIGISTKANSLIISWHSKSGNVERYLMLMDKGLVHGGVVDKHTSYAF 475
258 HGLSPGYLYNLVTWTEAAGLQNYRWKLVRTAPMEVSNLKVNTDGLSTLSLKKWQRPQGV 317
476 HGLTPGYLYNLVTWTEAAGLQNYRWKLVRTAPMEVSNLKVNTDGLSTLSLKKWQRPQGV 535
318 DSYNITLSHGKTIKESRVLAPWITETHFKELVPCRLYQVTVSCVSGELSAQKMAVGRTFP 377
536 DSYNITLSHGKTIKESRVLAPWITETHFKELVPCRLYQVTVSCVSGELSAQKMAVGRT-- 593
378 DKVANLEANNNGRMRSVLVSWSP 400
594 -----CESWAP 599

RESULT 13
AD004582
ID AD004582 standard; protein; 336 AA.
XX
AC AD004582;
XX
DT 15-JUL-2004 (first entry)
XX
DE Human HPTPbeta intracellular domain (ICD) protein.
XX
KW Protein co-ordinate data; HPTPbeta; HPTP-beta; PTPRB; PTPbeta; PTPB;
KW R-PTP-beta; angiogenesis mediated disorder; diabetic retinopathy;
KW sickle cell anaemia; Paget's disease; mycobacterial infection;
KW systemic lupus erythematosus; myopia; Crohn's disease; psoriasis;
KW rheumatoid arthritis; tumour; acquired immune deficiency syndrome; AIDS;
KW drug designing; therapy; human; intracellular domain; ICD.
XX
OS Homo sapiens.
XX
PN US2004077065-A1.
XX
PD 22-APR-2004.
XX
PF 04-AUG-2003; 2003US-00634027.
XX
PR 25-SEP-2002; 2002US-0413547P.
XX
PA (PROC ) PROCTER & GAMBLE CO.
XX
PI Evdokimov AG, Pokross ME;
XX
WPI; 2004-374235/35.
DR N-PSDB; AD004581.
XX
XX
Identification of compound useful for treatment of angiogenesis mediated
PT disorder, by using three-dimensional structure of HPTPbeta catalytic
PT domain, and employing structure to design, or select compound that binds
PT HPTPbeta in silico.
XX
XX
Disclosure; SEQ ID NO 4; 335pp; English.
XX
XX
The invention relates to the three dimensional coordinates of HPTPbeta
CC (also known as HPTP-beta, PTPRB, PTPbeta, PTPB or R-PTP-beta) protein. It
CC also relates to a method for the identification of a compound useful for
CC the treatment of an angiogenesis mediated disorder. The compounds
CC identified by this method are useful to treat diseases like diabetic
CC retinopathy, sickle cell anaemia, Paget's disease, mycobacterial
CC infections, systemic lupus erythematosus, myopia, Crohn's disease,
CC psoriasis, rheumatoid arthritis, solid or blood borne tumours and
CC acquired immune deficiency syndrome (AIDS). The invention is useful for
CC the treatment of an angiogenesis mediated disorder or disease. It is also
CC useful in drug design techniques. The present sequence is human HPTPbeta
CC intracellular domain (ICD) protein.
XX
SQ Sequence 336 AA;

```

Query Match 17.5%; Score 1830; DB 8; Length 336;  
Best Local Similarity 100.0%; Pred. No. 2.7e-120; Mismatches 0; Indels 0; Gaps 0;  
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1662 DRPLSVHLNLGQGNKRTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDI 1721  
DB 1 DRPLSVHLNLGQGNKRTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDI 60

QY 1722 ALLPENRGKRYNNILPYDATRVKLSNVDDDDPCSDYINASYIPGNFRREYIVTQGPLPG 1781  
DB 61 ALLPENRGKRYNNILPYDATRVKLSNVDDDDPCSDYINASYIPGNFRREYIVTQGPLPG 120

QY 1782 TKDDFWKMWVEQNVHNIWMVTQCVKGRVKCDHYWPAQDQSLXYGDLILQMLSESVLPEW 1841  
DB 121 TKDDFWKMWVEQNVHNIWMVTQCVKGRVKCDHYWPAQDQSLXYGDLILQMLSESVLPEW 180

QY 1842 TIREFKICGEEQLDAHLRIHFHYTVPDGHGVPETTSQSLIQFVRTVRYDINRSPGAGPTV 1901  
DB 181 TIREFKICGEEQLDAHLRIHFHYTVPDGHGVPETTSQSLIQFVRTVRYDINRSPGAGPTV 240

QY 1902 VHCAGVGRGTGTFIALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQTECOYVYLHOCVR 1961  
DB 241 VHCAGVGRGTGTFIALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQTECOYVYLHOCVR 300

QY 1962 DVLARKLRSEQENPLFPPIYENVNPEYHRDPVYSRH 1997  
DB 301 DVLARKLRSEQENPLFPPIYENVNPEYHRDPVYSRH 336

## RESULT 14

AD004585  
ID AD004585 standard; protein; 319 AA.

XX AC AD004585;

XX DT 15-JUL-2004 (first entry)

XX DE Human HPTPbeta catalytic domain.

XX KW Protein co-ordinate data; HPTPbeta; HPTP-beta; PTPRB; PTPbeta; PTPB;  
KW R-PTP-beta; angiogenesis mediated disorder; diabetic retinopathy;  
KW sickle cell anaemia; Paget's disease; mycobacterial infection;  
KW systemic lupus erythematosus; myopia; Crohn's disease; psoriasis;  
KW rheumatoid arthritis; tumour; acquired immune deficiency syndrome; AIDS;  
KW drug designing; therapy; human.

XX OS Homo sapiens.

XX PN US2004077065-A1.

XX PD 22-APR-2004.

XX PF 04-AUG-2003; 2003US-00634027.

XX PR 25-SEP-2002; 2002US-0413547P.

XX PA (PROC ) PROCTER & GAMBLE CO.

XX PI Evdokimov AG, Pokross ME;

XX DR WPI; 2004-374235/35.

XX PT Identification of compound useful for treatment of angiogenesis mediated disorder, by using three-dimensional structure of HPTPbeta catalytic domain, and employing structure to design, or select compound that binds HPTPbeta in silico.

XX PS Example; SEQ ID NO 7; 335pp; English.

XX CC The invention relates to the three dimensional coordinates of HPTPbeta (also known as HPTP-beta, PTPRB, PTPbeta, PTPB or R-PTP-beta) protein. It also relates to a method for the identification of a compound useful for the treatment of an angiogenesis mediated disorder. The compounds

CC identified by this method are useful to treat diseases like diabetic retinopathy, sickle cell anaemia, Paget's disease, mycobacterial infections, systemic lupus erythematosus, myopia, Crohn's disease, psoriasis, rheumatoid arthritis, solid or blood borne tumours and acquired immune deficiency syndrome (AIDS). The invention is useful for the treatment of an angiogenesis mediated disorder or disease. It is also useful in drug design techniques. The present sequence is human HPTPbeta protein catalytic domain.

XX SQ Sequence 319 AA;

Query Match 16.1%; Score 1691; DB 8; Length 319;

Best Local Similarity 99.4%; Pred. No. 1.7e-110;

Matches 312; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1662 DRPLSVHLNLGQGNKRTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDI 1721

DB 2 DRPLSVHLNLGQGNKRTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDI 61

QY 1722 ALLPENRGKRYNNILPYDATRVKLSNVDDDDPCSDYINASYIPGNFRREYIVTQGPLPG 1781

DB 62 ALLPENRGKRYNNILPYDATRVKLSNVDDDDPCSDYINASYIPGNFRREYIVTQGPLPG 121

QY 1782 TKDDFWKMWVEQNVHNIWMVTQCVKGRVKCDHYWPAQDQSLXYGDLILQMLSESVLPEW 1841

DB 122 TKDDFWKMWVEQNVHNIWMVTQCVKGRVKCDHYWPAQDQSLXYGDLILQMLSESVLPEW 181

QY 1842 TIREFKICGEEQLDAHLRIHFHYTVPDGHGVPETTSQSLIQFVRTVRYDINRSPGAGPTV 1901

DB 182 TIREFKICGEEQLDAHLRIHFHYTVPDGHGVPETTSQSLIQFVRTVRYDINRSPGAGPTV 241

QY 1902 VHCAGVGRGTGTFIALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQTECOYVYLHOCVR 1961

DB 242 VHCAGVGRGTGTFIALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQTECOYVYLHOCVR 301

QY 1962 DVLARKLRSEQEN 1975

DB 302 DVLARKLRSEQEN 315

## RESULT 15

AD004584

ID AD004584 standard; protein; 312 AA.

XX AC AD004584;

XX DT 15-JUL-2004 (first entry)

XX DE Human HPTPbeta truncated intracellular domain (ICD) protein.

XX KW Protein co-ordinate data; HPTPbeta; HPTP-beta; PTPRB; PTPbeta; PTPB;  
KW R-PTP-beta; angiogenesis mediated disorder; diabetic retinopathy;  
KW sickle cell anaemia; Paget's disease; mycobacterial infection;  
KW systemic lupus erythematosus; myopia; Crohn's disease; psoriasis;  
KW rheumatoid arthritis; tumour; acquired immune deficiency syndrome; AIDS;  
KW drug designing; therapy; human; intracellular domain; ICD; mutant; mutein.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN US2004077065-A1.

XX PD 22-APR-2004.

XX PF 04-AUG-2003; 2003US-00634027.

XX PR 25-SEP-2002; 2002US-0413547P.

XX PA (PROC ) PROCTER & GAMBLE CO.

XX PI Evdokimov AG, Pokross ME;

XX XX

DR WPI; 2004-374235/35.  
DR N-PSDB; ADO04583.  
XX  
XX Identification of compound useful for treatment of angiogenesis mediated  
PT disorder, by using three-dimensional structure of HTPbeta catalytic  
PT domain, and employing structure to design, or select compound that binds  
XX HTPbeta in silico.  
XX  
PS Disclosure; SEQ ID NO 6; 335pp; English.  
XX  
XX The invention relates to the three dimensional coordinates of HTPbeta  
CC (also known as HTP-beta, FPRB, FPRB, FPRB or R-PTP-beta) protein. It  
CC also relates to a method for the identification of a compound useful for  
CC the treatment of an angiogenesis mediated disorder. The compounds  
CC identified by this method are useful to treat diseases like diabetic  
CC retinopathy, sickle cell anaemia, Paget's disease, mycobacterial  
CC infections, systemic lupus erythematosus, myopia, Crohn's disease,  
CC psoriasis, rheumatoid arthritis, solid or blood borne tumours and  
CC acquired immune deficiency syndrome (AIDS). The invention is useful for  
CC the treatment of an angiogenesis mediated disorder or disease. It is also  
CC useful in drug design techniques. The present sequence is human HTPbeta  
CC intracellular domain (ICD) truncated protein.  
XX  
XX Sequence 312 AA;  
SQ  
Query Match 16.1%; Score 1690; DB 8; Length 312;  
Best Local Similarity 100.0%; Pred. No. 1.9e-110;  
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1662 DRPLSVHLNLGQGNKRTSCPIKINQFEGHPMKLQADSNYLLSKEYBELKDVGRNQSCDI 1721  
Db 1 DRPLSVHLNLGQGNKRTSCPIKINQFEGHPMKLQADSNYLLSKEYBELKDVGRNQSCDI 60  
Qy 1722 ALLPENRGNKNNILPYDATRVKLSNVDDPCSDYINASYIPGNFRREYIVTQGPLPG 1781  
Db 61 ALLPENRGNKNNILPYDATRVKLSNVDDPCSDYINASYIPGNFRREYIVTQGPLPG 120  
Qy 1782 TKDDFWKMWVQNVHNIWMVTQCVEKGRVKCDHYWPADQDSLYYGDLLQMLSESVLPEW 1841  
Db 121 TKDDFWKMWVQNVHNIWMVTQCVEKGRVKCDHYWPADQDSLYYGDLLQMLSESVLPEW 180  
Qy 1842 TIREFKICGEEQLDAHLRIHFHYTVPDHCVPETTSQSLIQFVTVRDYINRSPGAGTV 1901  
Db 181 TIREFKICGEEQLDAHLRIHFHYTVPDHCVPETTSQSLIQFVTVRDYINRSPGAGTV 240  
Qy 1902 VHCSAGVGRGTGTFALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQTECOYVYLHOCVR 1961  
Db 241 VHCSAGVGRGTGTFALDRILQQLDSKDSVDIYGAVHDLRLHRVHMVQTECOYVYLHOCVR 300  
Qy 1962 DVLRAKLRSEQ 1973  
Db 301 DVLRAKLRSEQ 312

Search completed: March 10, 2006, 18:26:36  
Job time : 223.124 secs

## Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.

- If you encounter an accession number from an older search run against UniProt (results file extension .rnp) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

**This Page Blank (uspio)**



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model  
Run on: March 10, 2006, 18:19:33 ; Search time 285.648 Seconds  
(without alignments)  
4932.432 Million cell updates/sec

Title: US-10-633-742-2  
Perfect score: 10483  
Sequence: 1 MLSHGAGLWTLISLLQTG.....PPIYENVPEVHRDPVYSRH 1997

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_05\_80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10483	100.0	1997	1 PTPRB_HUMAN	P23467 homo sapien
2	10226	97.5	1956	2 Q6MZFC_HUMAN	Q6mzf6 homo sapien
3	9124.5	87.0	1998	2 Q8CIW2_MOUSE	Q8ciw2 mus musculu
4	6712	64.0	1407	2 Q6ZRI9_HUMAN	Q6zri9 homo sapien
5	6703	63.9	1527	2 Q6AWC4_HUMAN	Q6awc4 homo sapien
6	5414	51.6	1102	2 Q8QVW7_MOUSE	Q8qv77 mus musculu
7	3863.5	36.9	1991	2 Q4TC72_TETNG	Q4tc72 tetraodon n
8	3413	32.6	919	2 Q88D73_HUMAN	Q88d73 homo sapien
9	2937	28.0	579	2 Q9JJO7_MOUSE	Q9jj07 mus musculu
10	2772	26.4	771	2 Q86VA4_HUMAN	Q86va4 homo sapien
11	1854.5	17.7	609	2 Q6ZTX7_HUMAN	Q6ztx7 homo sapien
12	1718.5	16.4	723	2 Q9CX77_MOUSE	Q9cx77 mus musculu
13	1566	14.9	1506	2 Q7PQLO_ANOGA	Q7pqlo anopheles g
14	1561.5	14.9	1631	1 PTP10_DROME	P35992 drosophila
15	1485.5	14.2	1607	2 Q8IRSO_DROSOPHILA	Q8irs0 drosophila
16	1485.5	14.2	1767	2 Q9W4F5_DROME	Q9w4f5 drosophila
17	1457.5	13.9	1767	2 Q24495_DROME	Q24495 drosophila
18	1379.5	13.2	1406	2 Q9W6V5_CHICK	Q9w6v5 gallus gall
19	1243.5	11.9	1238	1 PTPRJ_MOUSE	Q64455 m receptor-
20	1243.5	11.9	1238	2 Q841R5_MOUSE	Q841r5 mus musculu
21	1242.5	11.9	1238	2 Q8K3Q2_MOUSE	Q8k3q2 mus musculu
22	1242	11.8	1337	1 PTPRJ_HUMAN	Q12913 h receptor-
23	1233.5	11.8	1064	2 Q8XBX2_DROME	Q8xbx2 drosophila
24	1228.5	11.7	1238	2 Q8CIW9_MOUSE	Q8ciw9 mus musculu
25	1221.5	11.7	1216	2 Q62884_RAT	Q62884 rattus norv
26	1204	11.5	1705	2 Q8ERK5_MOUSE	Q8erk5 mus musculu
27	1200	11.4	1705	1 PTPRV_MOUSE	P70289 mus musculu
28	1197	11.4	1711	1 PTPRV_RAT	Q64612 rattus norv
29	1103.5	10.5	1188	2 Q4SDV5_TETNG	Q4sdy5 tetraodon n
30	1097	10.5	2302	2 Q88488_RAT	Q88488 rattus norv
31	1003	9.6	442	2 Q4RBH9_TETNG	Q4rbh9 tetraodon n

#### RESULT 1

ID	PTPRB_HUMAN	STANDARD;	PRT;	1997 AA.
AC	P23467;			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-NOV-1991 (Rel. 20, Last sequence update)			
DT	10-MAY-2005 (Rel. 47, Last annotation update)			
DE	Receptor-type tyrosine-protein phosphatase beta precursor (EC 3.1.3.48) (Protein-tyrosine phosphatase beta) (R-PTP-beta).			
GN	Name=PTPRB; Synonyms=PTPB;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;			
OC	Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	TISSUE=Placenta;			
RX	MEDLINE=91006018; PubMed=2170109;			
RA	Krueger N.X., Streuli M., Saito H.;			
RT	"Structural diversity and evolution of human receptor-like protein tyrosine phosphatases."			
RL	EMBO J. 9:3241-3252(1990).			
CC	-1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein tyrosine + phosphate.			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.			
CC	Receptor class 3 subfamily.			
CC	-1- SIMILARITY: Contains 17 fibronectin type-III domains.			
CC	-1- SIMILARITY: Contains 1 tyrosine-protein phosphatase domain.			
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.			
CC	EMBL; X54131; CAA38066.1; -; mRNA.			
DR	PIR; S12050; S12050.			
DR	HSSP; P18052; 1YFO.			
DR	Ensembl; ENSG00000127329; Homo sapiens.			
DR	HGNC; HGNC:9665; PTPRB.			
DR	MIM; 176882; -.			
DR	GO; GO:0005887; C:integral to plasma membrane; TAS.			
DR	GO; GO:0005001; P:transmembrane receptor protein tyrosine phosphatase.			
DR	GO; GO:0006796; P:phosphate metabolism; TAS.			
DR	GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.			
DR	InterPro; IPR003961; FN_III.			
DR	InterPro; IPR000387; TYR_phosphatase.			
DR	InterPro; IPR000242; Tyr_PP.			
DR	Pfam; PF00041; fn3; 16.			
DR	Pfam; PF00102; Y_phosphatase; 1.			
DR	PRINTS; PR00700; PTPRHPTASE.			
DR	SMART; SM00060; FN3; 17.			

#### ALIGNMENTS

ID	PTPRB_HUMAN	STANDARD;	PRT;	1997 AA.
AC	P23467;			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-NOV-1991 (Rel. 20, Last sequence update)			
DT	10-MAY-2005 (Rel. 47, Last annotation update)			
DE	Receptor-type tyrosine-protein phosphatase beta precursor (EC 3.1.3.48) (Protein-tyrosine phosphatase beta) (R-PTP-beta).			
GN	Name=PTPRB; Synonyms=PTPB;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;			
OC	Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	TISSUE=Placenta;			
RX	MEDLINE=91006018; PubMed=2170109;			
RA	Krueger N.X., Streuli M., Saito H.;			
RT	"Structural diversity and evolution of human receptor-like protein tyrosine phosphatases."			
RL	EMBO J. 9:3241-3252(1990).			
CC	-1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein tyrosine + phosphate.			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-1- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.			
CC	Receptor class 3 subfamily.			
CC	-1- SIMILARITY: Contains 17 fibronectin type-III domains.			
CC	-1- SIMILARITY: Contains 1 tyrosine-protein phosphatase domain.			
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.			
CC	EMBL; X54131; CAA38066.1; -; mRNA.			
DR	PIR; S12050; S12050.			
DR	HSSP; P18052; 1YFO.			
DR	Ensembl; ENSG00000127329; Homo sapiens.			
DR	HGNC; HGNC:9665; PTPRB.			
DR	MIM; 176882; -.			
DR	GO; GO:0005887; C:integral to plasma membrane; TAS.			
DR	GO; GO:0005001; P:transmembrane receptor protein tyrosine phosphatase.			
DR	GO; GO:0006796; P:phosphate metabolism; TAS.			
DR	GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.			
DR	InterPro; IPR003961; FN_III.			
DR	InterPro; IPR000387; TYR_phosphatase.			
DR	InterPro; IPR000242; Tyr_PP.			
DR	Pfam; PF00041; fn3; 16.			
DR	Pfam; PF00102; Y_phosphatase; 1.			
DR	PRINTS; PR00700; PTPRHPTASE.			
DR	SMART; SM00060; FN3; 17.			

DR	SMART; SM00194; PTPC; 1.	121	EKTTSTGLHVMWTPSSGKVTSEYQLFDENNQKIQGVQIQESTSWNEYTFNLTAGSKYN	180
DR	PROSITE; PS00853; FN3; 17.	121		
DR	PROSITE; PS00383; TVR_PHOSPHATASE_1; 1.	121	EKTTSTGLHVMWTPSSGKVTSEYQLFDENNQKIQGVQIQESTSWNEYTFNLTAGSKYN	180
DR	PROSITE; PS00056; TVR_PHOSPHATASE_2; 1.			
DR	PROSITE; PS00055; TVR_PHOSPHATASE_PTP; 1.	181	IAITAVSGGRRSPSVYTNSTGTPSPVKDIGISTKANSLLSWSHGSGNVERYRLMLMDKG	240
KW	Glycoprotein; Hydrolase; Protein phosphatase; Repeat;	181	IAITAVSGGRRSPSVYTNSTGTPSPVKDIGISTKANSLLSWSHGSGNVERYRLMLMDKG	240
KW	Signal; Transmembrane.			
FT	SIGNAL	1	22	Potential.
FT	CHAIN	23	1997	Receptor-type tyrosine-protein phosphatase beta.
FT				Extracellular (Potential).
FT	TOPO_DOM	23	1621	Potential.
FT	TRANSMEM	1622	1642	Cytoplasmic (Potential).
FT	TOPO_DOM	1643	1997	Fibronectin type-III 1.
FT	DOMAIN	23	108	Fibronectin type-III 2.
FT	DOMAIN	110	201	Fibronectin type-III 3.
FT	DOMAIN	203	288	Fibronectin type-III 4.
FT	DOMAIN	289	375	Fibronectin type-III 5.
FT	DOMAIN	377	465	Fibronectin type-III 6.
FT	DOMAIN	467	552	Fibronectin type-III 7.
FT	DOMAIN	554	641	Fibronectin type-III 8.
FT	DOMAIN	643	730	Fibronectin type-III 9.
FT	DOMAIN	731	817	Fibronectin type-III 10.
FT	DOMAIN	819	906	Fibronectin type-III 11.
FT	DOMAIN	907	993	Fibronectin type-III 12.
FT	DOMAIN	995	1083	Fibronectin type-III 13.
FT	DOMAIN	1085	1172	Fibronectin type-III 14.
FT	DOMAIN	1173	1260	Fibronectin type-III 15.
FT	DOMAIN	1261	1353	Fibronectin type-III 16.
FT	DOMAIN	1355	1443	Fibronectin type-III 17.
FT	DOMAIN	1448	1550	Tyrosine-protein phosphatase.
FT	DOMAIN	1703	1963	Phosphocysteine intermediate (By similarity).
FT	ACT_SITE	1904	1904	
FT		28	28	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	53	53	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	75	75	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	172	172	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	198	198	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	267	267	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	321	321	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	414	414	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	421	421	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	479	479	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	544	544	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	574	574	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	598	598	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	652	652	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	721	721	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	829	829	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1040	1040	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1096	1096	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1163	1163	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1185	1185	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1212	1212	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1274	1274	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1367	1367	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1470	1470	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1474	1474	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1518	1518	N-linked (GlcNAc. . .) (Potential).
SQ	SEQUENCE	1997	AA; 224268 MW; 691E99BA7A1515DD CRC64;	
	Query Match		100.0%; Score 10483; DB 1; Length 1997;	
	Best Local Similarity	100.0%;	Pred. No. 0;	
	Matches 1997; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;	
Qy	1	MLSHGAGLAWITLSLLQTGLAEPCNFTLAESKASHSVSIQWRILGSPCNFLIYSS	60	
Db	1	MLSHGAGLAWITLSLLQTGLAEPCNFTLAESKASHSVSIQWRILGSPCNFLIYSS	60	
Qy	61	DTLGAALCPTFRIDNTTYGCMQLDQAGTIYNFKIISLDEERTVVLQTDPLPPARFGVSK	120	
Db	61	DTLGAALCPTFRIDNTTYGCMQLDQAGTIYNFKIISLDEERTVVLQTDPLPPARFGVSK	120	

Db	1201	FDFFYELL	YNPNGT	KKENWKD	OLTEWR	FQGL	VPGR	KYVL	VVWV	THSGD	LSN	KYTAG	SRTA	1260			
Qy	1261	PSPPS	LSMFAD	IANTS	LAIT	TKWGP	PDWT	YND	PELO	MLPR	DALT	VFNPN	YNNR	KSEGR	IYV	1320	
Db	1261	PSPPS	LSMFAD	IANTS	LAIT	TKWGP	PDWT	YND	PELO	MLPR	DALT	VFNPN	YNNR	KSEGR	IYV	1320	
Qy	1321	GLRPRS	YQFN	KTVSG	DSWK	YTSK	PIFG	SVRT	KPKD	KIQL	NHCR	PQNS	TAIA	CSW	IPPD	1380	
Db	1321	GLRPRS	YQFN	KTVSG	DSWK	YTSK	PIFG	SVRT	KPKD	KIQL	NHCR	PQNS	TAIA	CSW	IPPD	1380	
Qy	1381	DFDGS	YTECR	KMDTQ	VEF	SR	KL	EK	ES	LI	NM	VL	VP	HK	YLV	SI	1440
Db	1381	DFDGS	YTECR	KMDTQ	VEF	SR	KL	EK	ES	LI	NM	VL	VP	HK	YLV	SI	1440
Qy	1441	TITMI	DR	PPPP	PH	IR	N	E	K	D	V	L	S	K	S	I	1500
Db	1441	TITMI	DR	PPPP	PH	IR	N	E	K	D	V	L	S	K	S	I	1500
Qy	1501	LKPEQ	QHPL	PS	LYE	R	N	A	S	I	R	V	Q	T	N	Y	1560
Db	1501	LKPEQ	QHPL	PS	LYE	R	N	A	S	I	R	V	Q	T	N	Y	1560
Qy	1561	PTQKF	CDG	PL	K	P	H	T	A	R	I	S	I	R	A	P	1620
Db	1561	PTQKF	CDG	PL	K	P	H	T	A	R	I	S	I	R	A	P	1620
Qy	1621	EGVS	AG	FL	T	G	M	L	V	A	V	A	L	I	C	R	1680
Db	1621	EGVS	AG	FL	T	G	M	L	V	A	V	A	L	I	C	R	1680
Qy	1681	CPT	K	I	N	O	F	E	G	H	F	M	K	L	O	A	1740
Db	1681	CPT	K	I	N	O	F	E	G	H	F	M	K	L	O	A	1740
Qy	1741	A	T	R	V	K	L	S	N	V	D	D	P	C	S	D	1800
Db	1741	A	T	R	V	K	L	S	N	V	D	D	P	C	S	D	1800
Qy	1801	V	T	O	C	V	E	G	R	V	K	D	H	Y	P	A	1860
Db	1801	V	T	O	C	V	E	G	R	V	K	D	H	Y	P	A	1860
Qy	1861	R	H	E	H	Y	T	V	M	P	D	H	G	V	P	E	1920
Db	1861	R	H	E	H	Y	T	V	M	P	D	H	G	V	P	E	1920
Qy	1921	L	Q	L	D	S	K	D	S	V	D	I	Y	G	A	V	1980
Db	1921	L	Q	L	D	S	K	D	S	V	D	I	Y	G	A	V	1980
Qy	1981	Y	E	N	V	N	P	E	H	R	D	P	V	Y	S	R	1997
Db	1981	Y	E	N	V	N	P	E	H	R	D	P	V	Y	S	R	1997

```

RESULT 2
Q6MZFF6 HUMAN
ID Q6MZFF6 HUMAN PRELIMINARY; PRT; 1956 AA.
AC Q6MZFF6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686E13109 (fragment).
GN Name=DKFZp686E13109;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
NCBI_TaxID=9606;
[1]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Human fetal kidney;
RC

```

RG	The German Human cDNA Consortium;
RA	Lauber J., Bahr A., Mewes H.W., Weill B., Amid C., Osanger A., Fobo G.,
RA	Han M., Wienand S.;
RL	Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR	EMBL; BX648771; CAB246198.1; - ; mRNA.
DR	HSP; P18031; IKAV
DR	GO; GO:0016787; F:hydrolase activity; IEA.
DR	GO; GO:0004725; P:protein tyrosine phosphatase activity; IEA.
DR	GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR	InterPro; IPR003961; FN III.
DR	InterPro; IPR003595; PTPc motif.
DR	InterPro; IPR000387; TYR_phosphatase.
DR	InterPro; IPR000242; Tyr_PP.
DR	Pfam; PF00041; fn3; 16.
DR	Pfam; PF00102; Y_phosphatase; 1.
DR	PRINTS; PR00700; PRTRYPHPTASE.
DR	SMART; SM00060; FN3; 16.
DR	SMART; SM00194; PTPc; 1.
DR	SMART; SM00404; PTPc_motif; 1.
DR	PROSITE; PS00853; FN3; 15.
DR	PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR	PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR	PROSITE; PS00055; TYR_PHOSPHATASE_PTP; 1.
KW	Hydrolase, Hypothetical protein.
FT	NON_IDR 1
SQ	SEQUENCE 1956 AA; 219892 MW; 08EC5C8FAF17221E CRC64;

  

	Query Match	97.5%; Score 10226; DB 2; Length 1956;
	Best Local Similarity	99.5%; Pred. No. 0;
	Matches 1947; Conservative	3; Mismatches 6; Indels 0; Gaps 0;
Qy	42 SIOWRIILGSCPNFLSYSDTLGAALCPTFRIDNTTYGCNLDQLDQAGTINFKIISLDEE	101
Dd	1 SIOWRIILGSCPNFLSYSDTLGAALCPTFRIDNTTYGCNLDQLDQAGTINFKIISLDEE	60
Qy	102 RTVVLQTDPLPPARFGVSKEKTTSTGLHWMTSSCGKVTSYEVLFDENNKIQGVQIOE	161
Dd	61 RTVVLQTDPLPPARFGVSKEKTTSTSLHWMTSSCGKVTSYEVLFDENNKIQGVQIOE	120
Qy	162 STSWNEYTFNLTAGSKYNIAITAVSGGKRFSFYVTNGSTVPSPVKDIGISTKANSLIS	221
Dd	121 STSWNEYTFNLTAGSKYNIAITAVSGGKRFSFYVTNGSTVPSPVKDIGISTKANSLIS	180
Qy	222 WSHGSGNVRYRLMLMDKGILVHGGVVDDKHATSYPFHGLSPGYLYNLTVMTAAAGLNQR	281
Dd	181 WSHGSGNVRYRLMLMDKGILVHGGVVDDKHATSYPFHGLTPGYLYNLTVMTAAAGLNQR	240
Qy	282 WKLVRTAPMEVSNLKVTNDGSLTSLKKVQRPQGNVDNSYNIITLSHGKTIKESRVLPWIT	341
Dd	241 WKLVRTAPMEVSNLKVTNDGSLTSLKKVQRPQGNVDNSYNIITLSHGKTIKESRVLPWIT	300
Qy	342 ETHFKELVPGRLVQVTVSCVSGELSQAQMVGRTFPDKVANLEANNGRMRSVLVWSVPP	401
Dd	301 ETHFKELVPGRLVQVTVSCVSGELSQAQMVGRTFPDKVANLEANNGRMRSVLVWSVPP	360
Qy	402 AGDWEQYRIILLFNDSVLLNITVKEETQYVMDDTGLVPCRQYVEVEIVESGNLKNSERC	461
Dd	361 AGDWEQYRIILLFNDSVLLNITVKEETQYVMDDTGLVPCRQYVEVEIVESGNLKNSERC	420
Qy	462 QGRTPVLAQLRVKCHANETSLSIMMQTPVAEWEKIYIIISLADRDLLLIHKSLSKDAKEFT	521
Dd	421 QGRTPVLAQLRVKCHANETSLSIMMQTPVAEWEKIYIIISLADRDLLLIHKSLSKDAKEFT	480
Qy	522 FTDLVPCRKYMAVTSTSGDLKNSSSVKGRTPVAQVTDLHVANQGMTSSLFTTNWTOAQGD	581
Dd	481 FTDLVPCRKYMAVTSTSGDLKNSSSVKGRTPVAQVTDLHVANQGMTSSLFTTNWTOAQGD	540
Qy	582 VEFQVOLLIHENVVIKNESISSETSRYSFHSLSKGSGLYSVVVTVTSGGISRRQVVEGRT	641
Dd	541 VEFQVOLLIHENVVIKNESISSETSRYSFHSLSKGSGLYSVVVTVTSGGISRRQVVEGRT	600
Qy	642 VPSSVSGVTYNNSGRNDZLSVSWLVAPEGVDNYEVTLSDHGKGVQSILVIAKSVRRECSFSS	701

Db 601 VPSSVSGVTNNSGRNDYLSWLLAPGVDVNYEVLTHSHDGKVVQSLVIAKSVRECSFSS 660  
Qy 702 LTPRLYTVITTTTSGKYENHSFQERTVPDKVQGVSVNSARSDYLRLVSVVHATGDFDH 761  
Db 661 LTPRLYTVITTTTSGKYENHSFQERTVPDKVQGVSVNSARSDYLRLVSVVHATGDFDH 720  
Qy 762 YEVTIKKNKNNFIQTKSPKSENECFVQLVPGRLYSVTVTTKSGQYEAENGGRITPEP 821  
Db 721 YEVTIKKNKNNFIQTKSPKSENECFVQLVPGRLYSVTVTTKSGQYEAENGGRITPEP 780  
Qy 822 VKDILTLNRSTEDLHVTVWSGANGVDQYIEIQLLFNDMKVPPFHLVNTATBYRTSLTPG 881  
Db 781 VKDILTLNRSTEDLHVTVWSGANGVDQYIEIQLLFNDMKVPPFHLVNTATBYRTSLTPG 840  
Qy 882 ROYKILVLTISGDVQQAIFEGFTVPSAVKNIHISPNGATDSLTVNMTPGGDDVDSYTVS 941  
Db 841 ROYKILVLTISGDVQQAIFEGFTVPSAVKNIHISPNGATDSLTVNMTPGGDDVDSYTVS 900  
Qy 942 AFRHSQKVDSTQIPKHVFETHFRLEAGEQYQIMIASVSGSLKNQIINVGRTPVPSVQGV 1001  
Db 901 AFRHSQKVDSTQIPKHVFETHFRLEAGEQYQIMIASVSGSLKNQIINVGRTPVPSVQGV 960  
Qy 1002 IADNAYSYSILVSWQKAAAGVAERYDILLTENGILLRNTSEPAATTKQHKPEDLTPGKY 1061  
Db 961 IADNAYSYSILVSWQKAAAGVAERYDILLTENGILLRNTSEPAATTKQHKPEDLTPGKY 1020  
Qy 1062 KIQILTVSGGLFSEKAEQTEGRTVPAAVTDLRI TENSTRHLSFRWTASEGELSWNIFLYN 1121  
Db 1021 KIQILTVSGGLFSEKAEQTEGRTVPAAVTDLRI TENSTRHLSFRWTASEGELSWNIFLYN 1080  
Qy 1122 PDGNLQERAQVDPLVQSFQNLQGRMYKMWIVTHSGELSNESFIFGRTPVPSVSHLRG 1181  
Db 1081 PDGNLQERAQVDPLVQSFQNLQGRMYKMWIVTHSGELSNESFIFGRTPVPSVSHLRG 1140  
Qy 1182 SNRNTDLSLFWNWPASGDFDFYELILYNPNGTCKENWKODLFEWRFQGLVPGKRYVLW 1241  
Db 1141 SNRNTDLSLFWNWPASGDFDFYELILYNPNGTCKENWKODLFEWRFQGLVPGKRYVLW 1200  
Qy 1242 VVTHSGDLSNKTVAESRTAPSPSLMSFADIANSLAITWKGPDPDWDYNDFELQWLPRD 1301  
Db 1201 VVTHSGDLSNKTVAESRTAPSPSLMSFADIANSLAITWKGPDPDWDYNDFELQWLPRD 1260  
Qy 1302 ALTVFNPNYNNRKSGRIVYGLRPGRSYQFNKTVSGDSWTKYSKPIFGSVRTKPKDKIQLN 1361  
Db 1261 ALTVFNPNYNNRKSGRIVYGLRPGRSYQFNKTVSGDSWTKYSKPIFGSVRTKPKDKIQLN 1320  
Qy 1362 HCRPQNSTAIACSWIPDPSPDGYSEICRKMDTQEVFESRKLEKESLLNIMMLVPHKRY 1421  
Db 1321 HCRPQNSTAIACSWIPDPSPDGYSEICRKMDTQEVFESRKLEKESLLNIMMLVPHKRY 1380  
Qy 1422 LVSLKVSAGWTSSEVDESDTITMIDRPPPPPHIRVNEKDVLSKSSINFTVNCWSFSDT 1481  
Db 1381 LVSLKVSAGWTSSEVDESDTITMIDRPPPPPHIRVNEKDVLSKSSINFTVNCWSFSDT 1440  
Qy 1482 NGAVKYFTVVVREAGSDELKPEQHQPLPSVLEYRHNASIRVQNTYFASKCAENPNNS 1541  
Db 1441 NGAVKYFTVVVREAGSDELKPEQHQPLPSVLEYRHNASIRVQNTYFASKCAENPNNS 1500  
Qy 1542 KSFNIKLGABMESLGGKRDPTQKFCDPGLKPHPTAYRISIRAFQQLFDEDLKFTKPLYS 1601  
Db 1501 KSFNIKLGABMESLGGKRDPTQKFCDPGLKPHPTAYRISIRAFQQLFDEDLKFTKPLYS 1560  
Qy 1602 DTFESLPITTESEPLFCAGIEGVSAGFLIGWLAVAVALLICROKVSGRERPSARLSIRR 1661  
Db 1561 DTFESLPITTESEPLFCAGIEGVSAGFLIGWLAVAVALLICROKVSGRERPSARLSIRR 1620  
Qy 1662 DRPLSVHLNLGQKNRKTSCPIKINQFEGHFMKLOADSNYLLSKEYBELKDVGQRNQCSDI 1721  
Db 1621 DRPLSVHLNLGQKNRKTSCPIKINQFEGHFMKLOADSNYLLSKEYBELKDVGQRNQCSDI 1680  
Qy 1722 ALLPENRGNKRYNNILPYDATRVKLSNVDDPCSDYINASYIPGNPNFRYIVTVQGPLPG 1781  
Db 1681 ALLPENRGNKRYNNILPYDATRVKLSNVDDPCSDYINASYIPGNPNFRYIVTVQGPLPG 1740

Qy 1782 TKDDFWKVVWQNVHNIWVWTCQVEKGRVKCDHYWPAQDSLYYGDILLOMLSESVLPPEW 1841  
Db 1741 TKDDFWKVVWQNVHNIWVWTCQVEKGRVKCDHYWPAQDSLYYGDILLOMLSESVLPPEW 1800  
Qy 1842 TIREFKICGEBQDLAHLIRHFHYTVWPDHGVPTTQSLIQFVTVRDYINRSPGAGPTV 1901  
Db 1801 TIREFKICGEBQDLAHLIRHFHYTVWPDHGVPTTQSLIQFVTVRDYINRSPGAGPTV 1860  
Qy 1902 VHCAGVGRGTFTFALDRILLOQLDSDSDYIYGAHVHDLRLHRVHMVQTECOYVYLHQCVR 1961  
Db 1861 VHCAGVGRGTFTFALDRILLOQLDSDSDYIYGAHVHDLRLHRVHMVQTECOYVYLHQCVR 1920  
Qy 1962 DVLRLARKLRSEQENPLFPPIYENVNPEYHRDPVYSRH 1997  
Db 1921 DVLRLARKLRSEQENPLFPPIYENVNPEYHRDPVYSRH 1956  
RESULT 3  
Q8CIW2\_MOUSE  
ID Q8CIW2\_MOUSE PRELIMINARY; PRT; 1998 AA.  
AC Q8CIW2;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Vascular endothelial protein tyrosine phosphatase.  
GN Name=Ptpnb;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Swiss Webster;  
RX MEDLINE=2220060; PubMed=1234928; DOI=10.1093/emboj/cdf497;  
RA Nawroth R., Poell G., Ranft A., Klep S., Samulowitz U., Fachinger G.,  
RA Golding M., Shima D.T., Deutsch U., Vestweber D.;  
RT "VE-PTP and VE-cadherin ectodomains interact to facilitate regulation  
of phosphorylation and cell contacts.";  
RL EMBO J. 21:4885-4895(2002).  
CC -|- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
DR EMBL; AY077755; AAU75813.1; -; mRNA.  
DR HSP; P18052; IYFO.  
DR Ensembl; ENSMUSG0000020154; Mus musculus.  
DR MGI; MGI:97809; Ptpnb.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0016787; F:hydrolase activity; IEA.  
DR GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA.  
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.  
DR InterPro; IPR003961; FN III.  
DR InterPro; IPR003987; TYR\_phosphatase.  
DR InterPro; IPR000242; Tyr\_PP.  
DR Pfam; PF00041; fn3; 16.  
DR PRINTS; PR0102; Y\_phosphatase; 1.  
DR SMART; SM00060; FN3; 16.  
DR SMART; SM00194; PTEC; 1.  
DR PROSITE; PS00853; FN3; 16.  
DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
DR PROSITE; PS00056; TYR\_PHOSPHATASE\_2; 1.  
DR PROSITE; PS00055; TYR\_PHOSPHATASE\_ptp; 1.  
KW Hydrolase; Repeat; Transmembrane\_ptp;  
SQ SEQUENCE 1998 AA; 224474 MW; B37EA6501EF03E5D CRC64;

Query Match 87.0%; Score 9124.5; DB 2; Length 1998;  
Best Local Similarity 85.7%; Pred. No. 0;  
Matches 1713; Conservative 131; Mismatches 153; Indels 1; Gaps 1;  
Qy 1 MLSHGAGLAWITLSLLOTGLAEPKCNFTLASKSKASHSVSIQWRILGSPCHPSLIYSS 60  
Db 1 MLRHGALTALWITLSVVTGVAQVKCNFTLLESRVSSLSASIQWRTFASPCNPSLIYSS 60

QY 61 DTLGAALCPTFRIDNTTYGCLQADQAGTIYNFKIISLD-EERTVVLQDPLPBARQVS 119  
DB 61 DTSGPMWCHPIRIDNTFYGCPKDLQAGTVNFRIIVSLDGEESTLVVLQDPLPBARFEVN 120  
QY 120 KEKTTSTGLHWMTFSSGKVTSEYVQLFDENQKIQGVQIQUESTMNEYTFNLTAGSKY 179  
DB 121 REKTASTTQVRWTPSSGKVSWEVQLFDHNNKIQEVQVESITWQSYTFNLTEGNSY 180  
QY 180 NIATAVSGGKRSFVYNGSTVSPVKDGIISTYKANSLLISWSHGSGNVERFLMLMDK 239  
DB 181 KVAITAVSGKRSFPVYNGSTVSPVKDGIISPNPNSLLISWSRGSGNVEQYRLVMDK 240  
QY 240 GILVHGGVWDKHTASYAFHGLSPCYLNLVTMTAAAGLQNYRWKLVRTAPMEVSNLKVTN 299  
DB 241 GAIQVDINVDKRDTSYAFHELTGPHLYNLITVWASGLQNSRWKLVRTAPMEVSNLKVTN 300  
QY 300 DGSITSLSKVQWRPGNVDSTNYITLSHGKTIKESRVLAPMTITETFKELVPGRLYQVTVS 359  
DB 301 DGRITSLSLVKWKQPGDVDSVITLSHQGTIKESKTLAPVTTETQFKDLVPGRLYQVTVS 360  
QY 360 CVSGELSAQKMAVGRTPFDKVANLEANNNGMRSLVSVSWPPAGDWBOYRILLFNDNVVL 419  
DB 361 CISGELSAKSAAGRTVPEKVRNLVSYNEIWMKSFTVNTWTPPAGDWEHYRILVLFNBSLVL 420  
QY 420 LNTVKGEEQYVMDTGLVPCROYEVEIVESGNLKNSERCQGRTPVLAVLQLRVKHAN 479  
DB 421 LNTVKGEBETHYALDGLLELPGROYEVEIVESGNLKNSERCQGRTPVLAVLQLRVKHAN 480  
QY 480 ETSLSIMWOTPVAEWKEKYSIISLADRLILLHKSLSKDAKFTFTDLVPGRYMATVTSIS 539  
DB 481 ETSLSIGITWRAPLGEWKEKYSIISLADRLILLHKSLSKDAKFTFTDLVPGRYMATVTSIS 540  
QY 540 GDLKNSSSVKGRTPVPAQVTDLHVANQMGTSLSFTNWTQAGDVBFYQVLLIHENNVVKN 599  
DB 541 GDLKQSSSIKGRTPVPAQVTDLHVANQMGTSLSFTNWTQAGDVBFYQVLLIHENNVVKN 600  
QY 600 SISSETSRYSFHSLSKSSLSVWTVTSGGSISSQVVEGRTVPSSVSGVTVNNSGNDY 659  
DB 601 SVSSDTSRYSFRALKPGSLXSVWTVTSGGSISSQVVEGRTVPSSVSGVTVNNSGNDY 660  
QY 660 LSVSWLVPAGDVNDYVTLSDHGKVSQSLVIAKSVRECSFSLTPGRLYTVITTRSGKY 719  
DB 661 LSVSWLVPAGDVNDYVTLSDHGKVSQSLVIAKSVRECSFSLTPGRLYTVITTRSGKY 720  
QY 720 ENHSFQERTVPDKVQGVSVNSARSYLYRSWVHATGDFDHYEVTIKNKNRFTQTKSIP 779  
DB 721 ASHSFTEERTVPDKVQGVSVNSARSYLYRSWVHATGDFDHYEVTIKNKNRFTQTKTIP 780  
QY 780 KSENECFVQLVPGRLYSVTVTSGQYEAQNGRTIPBPVKDLILNRSTEDLHVTV 839  
DB 781 KSENECFVQLVPGRLYSVTVTSGQYEAQNGRTIPBPVKDLILNRSTEDLHVTV 840  
QY 840 SGANGDVDOYEIQLLNDMKVFPFPHLVNTATEYRFTSLTPGROYKILVLITISGDVQOSA 899  
DB 841 SRANGDVDOYEIQLLNDMKVFPFPHLVNTATEYRFTSLTPGROYKILVLITISGDVQOSA 900  
QY 900 FIEGFTVPSAVKNTIHSPNGATDLSLVNWTGPGGDVDSYTVSAFRHSQKVDSDOTIPKHVF 959  
DB 901 FIEGLPVPSTVKNTHISANGATDRLMTWSPGGGDVDSYTVSAFRQEKVDSDOTIPKHAS 960  
QY 960 EHTFRLLEAGBOYQIMIASVSGSLKQINNVVGRTPVSPASQGVQVIADNAYSSYSLIVSWOKA 1019  
DB 961 EHTFRLLEAGBOYQIMIASVSGSLKQINNVVGRTPVSPASQGVQVIADNAYSSYSLIVSWOKA 1020  
QY 1020 AGVAERVDILLTENGILLRNTSEPATYKQKFPDLTPGKYKYQILTVSGGLPESKAOT 1079  
DB 1021 LGVAERVDILLTENGILLRNTSEPATYKQKFPDLTPGKYKYQILTVSGGLPESKESQA 1080  
QY 1080 EGRTPVPAVTLRTENSTRHLSFRWTASEGELSGWYIFLNPNGNLQEAQVDPVQVSF 1139  
DB 1081 EGRTPVPAVTLRTENSTRHLSFRWTASEGELSGWYIFLNPDPRTQEAQVDPVQVSF 1140  
QY 1140 SFQNLQGRMYKMWIVTHSGELSNESFIFGRTPVSPASVSHLGRSNRNTTDSLWFWNSPASG 1199

DB 1141 SFQNLQGRMYKMWIVTHSGELSNESFIFGRTPVSPAAVNHLLKGSURNTTDSLWFWNSPASG 1200  
QY 1200 DDFYELILYNPNGTKEKWKDKDLTEWRPQGLVPGRYKYLWVVTTHSGDLSNKKVTASERT 1259  
DB 1201 DDFYELILYNPNGTKEKWKDKDLTEWRPQGLVPGRYKYLWVVTTHSGDLSNKKVTGSGRT 1260  
QY 1260 APSPPSLMSFADJANTSLAITWKGPDPDWDYNDYFQGLVPGRYKYLWVVTTHSGDLSNKKVTGSGRT 1319  
DB 1261 APSPPSLMSFADJANTSLAITWKGPDPDWDYNDYFQGLVPGRYKYLWVVTTHSGDLSNKKVTGSGRT 1320  
QY 1320 YGLRPGRSYFNKVTYSGDSWKTYSKPIFGSVRTKPKDKIQLNLCRPNQSTALACSWIPPD 1379  
DB 1321 YGLHPSYQFSVKTYSYSGDSWKTYSKPIFGSVRTKPKDKIQLNLCRPNQSTALACSWIPPD 1380  
QY 1380 SDFDGSYIECRKMDTQEVESRKLEKEKSLNLMMLVPHKRYLVSIIKQVSGAGTSEVVED 1439  
DB 1381 SDFDGSYIECRKMDTQEVESRKLEKEKSLNLMMLVPHKRYLVSIIKQVSGAGTSEVVED 1440  
QY 1440 STITMIDRPPPPHPIRVNEKDVLIISKSSINFTVNCWFSWFDSTNGAVKYPVVVREADGSD 1499  
DB 1441 STITMIDRPPPPHPIRVNEKDVLIISKSSINFTVNCWFSWFDSTNGAVKYPVVVREADGSD 1500  
QY 1500 ELKPEQOHPPLSPYLEVRHNASIRVYQNTYFASKCAENPNSNSKSFNIIKGAEMSLGCKR 1559  
DB 1501 ELKPEQOHPPLSPYLEVRHNASIRVYQNTYFASKCAENPNSNSKSFNIIKGAEMSLGCKR 1560  
QY 1560 DPTQKQFCDCGPLKPHPTAYRISIRAFQTLFDEDLKEFTKPLYSDDTFPSLPITTESSEPLFGA 1619  
DB 1561 DPTQKQFCDCGPLKPHPTAYRISIRAFQTLFDEDLKEFTKPLYSDDTFPSLPITTESSEPLFGV 1620  
QY 1620 IEGVSAGLFLIGMLVAVALLICROKVSQHSRERPSARLSIRDRPLSVHNLGQKGNRKT 1679  
DB 1621 IEGVSAGLFLIGMLVAVALLICROKVSQHSRERPSARLSIRDRPLSVHNLGQKGNRKT 1680  
QY 1680 SCPIKINQFEGHFMKQADSNYLLSKYBELKDVGRNQSCDIALLPENRGKRNRYNNILPY 1739  
DB 1681 SCPIKINQFEGHFMKQADSNYLLSKYBELKDVGRNQSCDIALLPENRGKRNRYNNILPY 1740  
QY 1740 DATRVKLSNVDDPCSDYINASYIPGNFRREYIVTQGLPGTKDDPFQKMWQNVNHNIV 1799  
DB 1741 DATRVKLSNVDDPCSDYINASYIPGNFRREYIVTQGLPGTKDDPFQKMWQNVNHNIV 1800  
QY 1800 MVTQCEKGRVKCDHWTPADQDSLYYGDLLIQLMSLVSEVWIREFKICGBEOLDARHL 1859  
DB 1801 MVTQCEKGRVKCDHWTPADQDSLYYGDLLIQLMSLVSEVWIREFKICGBEOLDARHL 1860  
QY 1860 IRHFHYTVWPDHGVPEPTQSLIQFVRTVTDYINRSYPCAGPTVHVHCSAGVGTGTGFIALDR 1919  
DB 1861 IRHFHYTVWPDHGVPEPTQSLIQFVRTVTDYINRSYPCAGPTVHVHCSAGVGTGTGFIALDR 1920  
QY 1920 ILQQLJSDKSDVYIGAVHDLRLHRVHMVQTECOVYVYLHQCVRDVLRAKLRSEQENPLFP 1979  
DB 1921 ILQQLJSDKSDVYIGAVHDLRLHRVHMVQTECOVYVYLHQCVRDVLRAKLRSEQENPLFP 1980  
QY 1980 IYENVNPEYHRDPIYSRH 1997  
DB 1981 IYENVNPEYHRDPIYSRH 1998

## RESULT 4

Q6ZR19 HUMAN

ID Q6ZR19 HUMAN PRELIMINARY; PRT; 1407 AA.

AC Q6ZR19;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE CDNA FLJ476721 fls, clone TRACH3018524, highly similar to Protein-

DE tyrosine phosphatase beta (EC 3.1.3.48). (Fragment).

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

```

OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Trachea;
RA Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S.,
RA Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Ota T., Wakamatsu A.,
RA Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshina A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK128562; BAC67502.1; -; mRNA.
DR GO; GO:0005529; F-sugar binding; IEA.
DR InterPro; IPR003961; FN_III.
DR Pfam; PF00041; fn3; 15.
DR SMART; SM00060; FN3; 14.
DR PROSITE; PS0853; FN3; 14.
FT NON TER 1407
SQ SEQUENCE 1407 AA; 156977 MW; 67B805B04FF50B65 CRC64;

Query Match 64.0%; Score 6712; DB 2; Length 1407;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1288; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 18 QTGLAEPERCNFTLAESKASHSVSIQWRILGSPCNFSLIYSSDTLGAALCPTFRIDNTT 77
Db 115 ETGLAEPERCNFTLAESKASHSVSIQWRILGSPCNFSLIYSSDTLGAALCPTFRIDNTT 174

Qy 78 YGCNLQDLQAGTIYFNRIISLDEBRTVLTQDPLPPARFGVSKETKTSTGLHVMWTPSSG 137
Db 175 YGCNLQDLQAGTIYFNRIISLDEBRTVLTQDPLPPARFGVSKETKTSTLHVMWTPSSG 234

Qy 138 KVTSYEVQLDENNQKIQGVQIOBSTSWNEYTFNLTAGSKYNIATAVSGGKRSPSVYT 197
Db 235 KVTSYEVQLDENNQKIQGVQIOBSTSWNEYTFNLTAGSKYNIATAVSGGKRSPSVYT 294

Qy 198 NGSTVPSFVKDIGISTKANSLLSWSHSGNVERYLMLMDKGLVHGGVVDKHTSYAF 257
Db 295 NGSTVPSFVKDIGISTKANSLLSWSHSGNVERYLMLMDKGLVHGGVVDKHTSYAF 354

Qy 258 HGLSPGYLYNLVTMEAGLQNYRWKLVRTAPMEVSNLKVTDNGSLTSLKVKWQRPFGNV 317
Db 355 HGLTPGYLYNLVTMEAGLQNYRWKLVRTAPMEVSNLKVTDNGSLTSLKVKWQRPFGNV 414

Qy 318 DSYNITLSHGKTIKESRVLPAPITETHPKELVPGRLYQVTVSCVSGELSAQKMAVGRTFP 377
Db 415 DSYNITLSHGKTIKESRVLPAPITETHPKELVPGRLYQVTVSCVSGELSAQKMAVGRTFP 474

Qy 378 DKVANLEANNNGMRSLVSVSWSPAGDQWQVRIILFNDSVVLNLTIVGKEETQVMDDTG 437
Db 475 DKVANLEANNNGMRSLVSVSWSPAGDQWQVRIILFNDSVVLNLTIVGKEETQVMDDTG 534

Qy 438 LVPGQRYEVEYIVESGNLKNSERCQGRTPVLAVLQLRVKHANETSLSIMQTPVAWEKY 497
Db 535 LVPGQRYEVEYIVESGNLKNSERCQGRTPVLAVLQLRVKHANETSLSIMQTPVAWEKY 594

Qy 498 IISLADRLLLIHKLSKDAKEFTFTDLVPGRYKMATVTSISGDLKNSSVKGRTPVAQV 557
Db 595 IISLADRLLLIHKLSKDAKEFTFTDLVPGRYKMATVTSISGDLKNSSVKGRTPVAQV 654

Qy 558 TDLHVNAGMTSSILFTNWTQAGDVEYQVLLITHENVVIKNESISSTSYSHLSKSGS 617
Db 655 TDLHVNAGMTSSILFTNWTQAGDVEYQVLLITHENVVIKNESISSTSYSHLSKSGS 714

Qy 618 LYSVWVTTVSGGISRRQVVEGRTVPSVSGVTVNNSGRNDYLSVSWLAPGVDVNYEVT 677
Db 715 LYSVWVTTVSGGISRRQVVEGRTVPSVSGVTVNNSGRNDYLSVSWLAPGVDVNYEVT 774

Qy 678 LSHDGKVVQSLVIATAKSVRECSFSLTPGRLYTVTITTRSGKYENHSPSQBRTVPDKQGV 737
Db 775 LSHDGKVVQSLVIATAKSVRECSFSLTPGRLYTVTITTRSGKYENHSPSQBRTVPDKQGV 834

```

## RESULT 5

Q6AWC4\_HUMAN

ID Q6AWC4\_HUMAN PRELIMINARY; PRT; 1527 AA.

AC Q6AWC4; 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Hypothetical protein DKF2p686E2262 (Fragment).

GN Name=DKF2p686E2262;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Endometrium;

RG The German cDNA Consortium;

RA Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,

RA Fobo G., Han M., Wiemann S.;

RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; BX47238; CAH10386.1; -; mRNA.

DR InterPro; IPR003961; FN\_III.

DR InterPro; IPR000772; Rigin\_B\_lectin.

DR Pfam; PF00041; fn3; 15.

DR SMART; SM00060; FN3; 14.

DR PROSITE; PS0853; FN3; 14.

DR PROSITE; PS0231; RIGIN\_B\_LECTIN; 1.

KW Hypothetical protein.

FT NON TER 1527

SQ	SEQUENCE	1527 AA; 170309 MW; 81A3DABACCO1BE6B CRC64;
	Query Match	63.9%; Score 6703; DB 2; Length 1527;
	Best Local Similarity	99.4%; Pred. No. 0;
	Matches 1285; Conservative	7; Mismatches 1; Indels 0; Gaps 0;
Qy	18 QTGLAEPERCNFTLAESKASHSVSIOWRILGSPCNFSLIYSSDTLGAALCPTFRIDNTT	77
Db	235 ETGLAEPERCNFTLAESKASHSVSIOWRILGSPCNFSLIYSSDTLGAALCPTFRIDNTT	294
Qy	78 YGCNLDQAGTIYNFKIISLDEERTVVLQDPLPPARFGVSKETTTSTGLHVMWTFSSG	137
Db	295 YGCNLDQAGTIYNFKIISLDEERTVVLQDPLPPARFGVSKETTTSTGLHVMWTFSSG	354
Qy	138 KVTSEYVQLFDENNOKIQGVQIQSTSWNEYTFNLTAGSKYNTAITAVSGGKRSFSVYT	197
Db	355 KVTSEYVQLFDENNOKIQGVQIQSTSWNEYTFNLTAGSKYNTAITAVSGGKRSFSVYT	414
Qy	198 NGSTVPSVVDIGISTKANSLLISWSHSGSNVERYRLMLMDKGLVHGVDVDDKATSYAF	257
Db	415 NGSTVPSVVDIGISTKANSLLISWSHSGSNVERYRLMLMDKGLVHGVDVDDKATSYAF	474
Qy	258 HGLSPGYLYNLVTMEAGLQNYRWKLVRTAPMEVSNLKVNTDGLSTSLKVKWQRPQGNV	317
Db	475 HGLTPGYLYNLVTMEAGLQNYRWKLVRTAPMEVSNLKVNTDGLSTSLKVKWQRPQGNV	534
Qy	318 DSYNITLSHKGTIKESRVLPAPWITETHFELKLPGRLYQVTVSCVSGELSAQKMAVGRTTP	377
Db	535 DSYNITLSHKGTIKESRVLPAPWITETHFELKLPGRLYQVTVSCVSGELSAQKMAVGRTTP	594
Qy	378 DKVANLEANNRMRSLVSVNSPAGDQWQVRIILLFNDVSVLLNITVKGKETOVMDDTG	437
Db	595 DKVANLEANNRMRSLVSVNSPAGDQWQVRIILLFNDVSVLLNITVKGKETOVMDDTG	654
Qy	438 LVPGQRYEVEVIVESGNLKNSERCQGRTPVPLAVLQVLRVGHANETSLSIMQTPVAWEKY	497
Db	655 LVPGQRYEVEVIVESGNLKNSERCQGRTPVPLAVLQVLRVGHANETSLSIMQTPVAWEKY	714
Qy	498 IISLADRLILLHKSLSKDAKEFTFDLPGRKYKMATVTSISGLDNSSSVKGRTPAQV	557
Db	715 IISLADRLILLHKSLSKDAKEFTFDLPGRKYKMATVTSISGLDNSSSVKGRTPAQV	774
Qy	558 TDLHVNQGTSSLTFTWNTQAGDVEFYQVLLIHENVVINKNESISSETSYSHSLKSGS	617
Db	775 TDLHVNQGTSSLTFTWNTQAGDVEFYQVLLIHENVVINKNESISSETSYSHSLKSGS	834
Qy	618 LYSVWVTVSGGISRQVWVEGRTPVSVSGVTWNSGRNDYLSVWLVAPGDVDNTEVT	677
Db	835 LYSVWVTVSGGISRQVWVEGRTPVSVSGVTWNSGRNDYLSVWLVAPGDVDNTEVT	894
Qy	678 LSHDGKVVQSLVIAKSVRECSFSLTPCRLTYTTITRSGKYENHSPSQERTVPDKQGV	737
Db	895 LSHDGKVVQSLVIAKSVRECSFSLTPCRLTYTTITRSGKYENHSPSQERTVPDKQGV	954
Qy	738 SVNSARSQVLYRWSVWVHATGDFHYETIKNNFIQTKSPKSENECVFQVLPGRLYS	797
Db	955 SVNSARSQVLYRWSVWVHATGDFHYETIKNNFIQTKSPKSENECVFQVLPGRLYS	1014
Qy	798 VTVTTKSGQYEAHQNGRTIPEPVKDLTLNRSTEDLHVWTSANGDQVQYEIQLLPND	857
Db	1015 VTVTTKSGQYEAHQNGRTIPEPVKDLTLNRSTEDLHVWTSANGDQVQYEIQLLPND	1074
Qy	858 MKVFPFPHLVNTATEYRFTSLTPGRQYKILVLTISGDVQQAFTLEGTFPSAVKNIHISP	917
Db	1075 MKVFPFPHLVNTATEYRFTSLTPGRQYKILVLTISGDVQQAFTLEGTFPSAVKNIHISP	1134
Qy	918 NGATDSLTVNWTGGGVDSVTSAPRHSQKVDSTQTPKHVFETHFRLRAGEQYQIMIA	977
Db	1135 NGATDSLTVNWTGGGVDSVTSAPRHSQKVDSTQTPKHVFETHFRLRAGEQYQIMIA	1194
Qy	978 SVSGSLKQKINVQGRTPVSPASQGVQVADNAYSSYSLIVSWQKAAVARYDILLTENGIL	1037
Db	1195 SVSGSLKQKINVQGRTPVSPASQGVQVADNAYSSYSLIVSWQKAAVARYDILLTENGIL	1254

Qy	1038 LRNTSEBATTQKHKFDLTGPKKYKIQIILTVSGLFSKAEQTEGRTVPAAVTDLRITENS	1097
Db	1255 LRNTSEBATTQKHKFDLTGPKKYKIQIILTVSGLFSKAEQTEGRTVPAAVTDLRITENS	1314
Qy	1098 TRHLSFWTASEGELSWYNIFLYNPDGNLQRAQVDPLVQSFQNLQGRMYKQVIVTH	1157
Db	1315 TRHLSFWTASEGELSWYNIFLYNPDGNLQRAQVDPLVQSFQNLQGRMYKQVIVTH	1374
Qy	1158 SGELSNESFIFGRTPVSPASVSHLRGSNNTTDSLWFWNSPAGDPDFYELILYNPNKTKCE	1217
Db	1375 SGELSNESFIFGRTPVSPASVSHLRGSNNTTDSLWFWNSPAGDPDFYELILYNPNKTKCE	1434
Qy	1218 NWKDKOLTEWFOGLVPGRKVYLVWVTHSGDLSNKNVTAESRTAPSPSLMSFADIANTSL	1277
Db	1435 NWKDEDTEWFOGLVPGRKVYLVWVTHSGDLSNKNVTAESRTAPSPSLMSFADIANTSL	1494
Qy	1278 AITWKGPPDMDVNDPELQWLPDALTIVFNPYN	1310
Db	1495 AITWKGPPDMDVNDPELQWLPDALTIVFNPYN	1527
RESULT 6		
Q80VN7_MOUSE PRELIMINARY; PRT; 1102 AA.		
ID	Q80VN7_MOUSE PRELIMINARY; PRT; 1102 AA.	
AC	Q80VN7; 01-JUN-2003 (TrEMBLrel. 24, Created)	
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)	
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)	
DE	Ptprb protein (Fragment).	
GN	Name=Ptprb;	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;	
OC	Muridae; Murinae; Mus.	
OX	NCBI_TaxID=10090;	
RN	[1]	
NUCLEOTIDE SEQUENCE.		
RP	STRAIN=NNRI; TISSUB=Mammary tumor. WAP-Tag model. 5 months old;	
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;	
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,	
RA	Klausner R.D., Collins F.S., Wagner L.H., Shenmen C.M., Schuler G.D.,	
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,	
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,	
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,	
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,	
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,	
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Harte S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Whiting M., Madan A., Kettelman M., Madan A., Rodriguez S., Sanchez A.,	
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,	
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,	
RA	Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smallus D.E.,	
RA	Schneerch A., Schein J.E., Jones S.J.M., Marra M.A.;	
RT	"Generation and initial analysis of more than 15,000 full-length human	
RT	and mouse cDNA sequences."	
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	
RN	[2]	
NUCLEOTIDE SEQUENCE.		
RP	STRAIN=NNRI; TISSUB=Mammary tumor. WAP-Tag model. 5 months old;	
RA	Strausberg R.;	
RL	Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; BC047086; AAH47086.1; -; mRNA.	
DR	HSSP; P18052; IYFO.	
DR	MGI; MGI:97809; Ptprb.	
DR	GO; GO:0016787; F:hydrolase activity; IEA.	
DR	GO; GO:0004725; P:protein tyrosine phosphatase activity; IEA.	
DR	GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.	
DR	InterPro; IPR003961; FN III.	
DR	InterPro; IPR000387; TYR_phosphatase.	



```

DR InterPro; IPR000242; Tyr_PP.
DR Pfam; PF00041; fn3; 6.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00700; TRYPHPTASE.
DR SMART; SM00060; FN3; 6.
DR SMART; SM00194; PTPC; 1.
DR PROSITE; PS00853; FN3; 6.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00556; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00555; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase.
FT NON TER
SQ SEQUENCE 1102 AA; 124677 MW; 8281DF7F29BF9148 CRC64;

Query Match 51.6%; Score 5414; DB 2; Length 1102;
Best Local Similarity 91.3%; Pred. No. 9e-307;
Matches 1006; Conservative 46; Mismatches 50; Indels 0; Gaps 0;

Qy 896 QQSAFIEGFTVPSAVKNIHISPGATDSLTVMWTPGGGDVDSYTSAFRHSQKVDSTIP 955
Db 1 QQSAFIEGLTVPSIVKNIHISANGATDRLVMWTPSGGGDVDSYVSAPRQDEKVDSTIP 60

Qy 956 KHVEHTFHRLEAGEQIQIMTASVSGSLKQINQVGRTPASVQGVADNAYSYSLIVS 1015
Db 61 KHASEHTFHRLEAGAKYRIATVSVSGSLRNQIDALGQTPASVQGVAAANAYSSNLTVS 120

Qy 1016 WQKAAGVAERYDILLTENGILLRNTSEPTATTKQHKFEDLTPGKKYKIQILTVSGGLFSK 1075
Db 121 WQKALGVAERYDILLNENGLLSNVSEPTARQHKFEDLTPGKKYKQILTVSGGLFSK 180

Qy 1076 EAQTEGRTVPAATDLRITENSTRHLSFRWTASEGELSWNIFLYPNPDGMLQERAAQVDPL 1135
Db 181 ESQAEGRTVPAATVNLRAITENSSRYLSFGWTASEGELSWNIFLYPNDRTLQERAAQVDPL 240

Qy 1136 VQSFSFQNLQGRMYKQVIVTHSGELSNESPIFGRTVPASVSHLGRGNRTTDSLWNWS 1195
Db 241 VQSFQNLQGRMYKQVIVTHSGELSNESPIFGRTVPAAVNLKGSHRNTTDSLWFSWS 300

Qy 1196 PASGDFDFYELILNPNGTCKENWKDLEWRFQGLVPGKRYLVWVTHSGDLSNKVTA 1255
Db 301 PASGDFDFYELILNPNGTCKENWKEDVTEWRFQGLVPGKRYLVVTHSGDLSNKVTG 360

Qy 1256 ESRTAPSPPSLMSPADIANSTLAITWKGPDPDWTYNDPELQWLPRDALTFNPNYNNRKE 1315
Db 361 EGRTAPSPSLLSPADVANTSLAITWKGPDPDWTYNDPELQWPFGDALTIFNPYSSRKSE 420

Qy 1316 GRIVYGLRPGRSYQFNKTVSGDSWKYTSKPIFGSVRTKPKDKIQLNLCRPNSTAIACSW 1375
Db 421 GRIVYGLRPGRSYQFSVKTVSGDSWKYTSKPIGSGVTRTKPKDKIQLNLCRPNSTAIACSW 480

Qy 1376 IPPDSDFGYSIECRKMDTOREVSFRKLEKEKSLNIMLVPHKRYLVSKVQSAGMTSE 1435
Db 481 IPPDSDFGYSIECRKMDTOREVSFRKLEKEKSLNIMLVPHKRYLVSKVQSAGMTSE 540

Qy 1436 VVEDSTITMDRPPPPPHIRVNEKQVLISKSSINFVNCWSFSDTNGAVKYFTVVVREA 1495
Db 541 VVEDSTITMDRPPPPPHIRVNEKQVLISKSSINFVNCWSFSDTNGAVKYFAVVVREA 600

Qy 1496 DGSDELKPEQHPPLPSYLEYRHNASIRVQNTYFASKCANPNNSNSFKIKLGAEMSL 1555
Db 601 DSMDELKPEQHPPLPSYLEYRHNASIRVQNTYFASKCAESPDSSSSFKIKLGAEMDSL 660

Qy 1556 GKGEDPTQKFCDCGKPLKPHYATYRISIRAFQLPDEDLKEFTKPLYSPTFFSLPTTSEEP 1615
Db 661 GKGCDPSQKFCDCGKPLKPHYATYRISIRAFQLPDEDLKEFTKPLYSPTFFSMPITTSEEP 720

Qy 1616 LFQAGIEGVSAGLFLIGMLVAVALLICRQKVSHERPERSARLSIRRRDPLSVHLNLGQKG 1675
Db 721 LFGVIEGVSAGLFLIGMLVAVAFVICRQKASHSRERSARLSIRRRDPLSVHLNLGQKG 780

Qy 1676 NRKTSCTPIKNQFEGHPMKLQADSNYLLSKEYBELKQVGRNQSCDIALLENRGNKRYNN 1735
Db 781 NRKTSCTPIKINQFEGHPMKLQADSNYLLSKEYEDLKQVGRSQSCDIALLENRGNKRYNN 840

```

```

Qy 1736 ILPYDATRVKLSNVDDPDCSDYINASYIPGNPNFREYIVTQGPLGTGKDDFWKMWQENV 1795
Db 841 ILPYDASRVKLCNVDDPDCSDYINASYIPGNPNFREYIATQGPLGTGKDDFWKMWQENV 900

Qy 1796 HNIWMVTQCEKGRVKCDHYWPAQDQSLXYGLDILQMLSESVLPETWIRFVKICSEBOLD 1855
Db 901 HNIWMVTQCEKGRVKCDHYWPAQDQPLXYGLDILQMVSESVLPETWIRFVKICSEBOLD 960

Qy 1856 ARLIRHFHYTVPDHCVPPTTQSLIQFVTRVDRDYNRSFGAGPTVVHCSAGVCRGTGTFI 1915
Db 961 ARLIRHFHYTVPDHCVPPTTQSLIQFVTRVDRDYNRSFGAGPTVVHCSAGVCRGTGTFV 1020

Qy 1916 ALDRILQOLDSKDSVDIYGAVHDLRLHRVHMVQTCOVYVYLHQCVRDVLRAKLRSQEN 1975
Db 1021 ALDRILQOLDSKDSVDIYGAVHDLRLHRVHMVQTCOVYVYLHQCVRDVLRAKLRSQEN 1080

Qy 1976 PLFPPIYENVNPEYHRDPVYSRH 1997
Db 1081 PLFPPIYENVNPEYHRDAIYSRH 1102

RESULT 7
Q4TC72_TETNG
ID Q4TC72 TETNG PRELIMINARY; PRT; 1991 AA.
AC Q4TC72;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAF7048, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00003440001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud C., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segreus B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Blemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAEE01007048; CAF89510.1; -; Genomic_DNA.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003595; PTPC motif.
DR InterPro; IPR000387; TYR_phosphatase.
DR InterPro; IPR000242; Tyr_PP.
DR Pfam; PF00041; fn3; 14.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00700; TRYPHPTASE.
DR SMART; SM00060; FN3; 14.
DR SMART; SM00194; PTPC; 1.
DR SMART; SM00404; PTPC_motif; 1.

```



DR PROSITE; PS50853; FN3; 14.  
DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
DR PROSITE; PS50056; TYR\_PHOSPHATASE\_2; 1.  
DR PROSITE; PS50055; TYR\_PHOSPHATASE\_PTP; 1.  
KW Hydrolase.  
FT NON\_TER 1 1991  
SQ SEQUENCE 1991 AA; 214280 MW; AE8CE4629DD27A78 CRC64;  
  
Query Match 36.9%; Score 3863.5; DB 2; Length 1991;  
Best Local Similarity 41.1%; Pred. No. 7.5e-216;  
Matches 831; Conservative 314; Mismatches 611; Indels 267; Gaps 29;  
  
QY 203 PSPVKDIGISKANSLI-LSWHSQGNVVERVRLMLMDKGIILVHGVDVKHATSAFGLS 261  
DB 1 PSAVGLVMTSRTSSLSGLSGAGPRTQRTQLQMLHRSALLQDQTQSTATQHTLLGLT 60  
  
QY 262 PGYLYNLVTMTAAAGLQWYRKLVRTAPMEYSNLKVTNDGSLTSLKVKWQPPGQNVDSYN 321  
DB 61 AGRYNYVTVEAGLRSSVTIEBQTPAAVSNITADASGG-TRVRLSWQRPQGDLDALV 119  
  
QY 322 ITLSHGKTIKESRVLAPWITHTFKELVPGRLYQVTVSCVSGELSAQMAVGRTPPKVA 381  
DB 120 VRLAANGDWMRATLPPNATEVWVDELPGWTYGVAVVSRGQRSSQSESSFGTAPAAAT 179  
  
QY 382 NL-----EANNNGMRSLVSWSPAGDWEQYRILLFNDSVVLINITVGKEETOYVMDT 436  
DB 180 RLILSPATAPGG-----LLLSWAPAGRWORYRLLRLLRDASLLVDAGVDOAVNYTFAGS 235  
  
QY 437 GLVPGROYEVVIVESGNLKNSECOGRT-----VPLAV 470  
DB 236 GLTPEGSLYAVLVESGGLTANSCEGATGVFLOGVTWAGSSINPLKGTSTTVSVVPAV 295  
  
QY 471 LQLRVGHANETSLIMW-QTPVAWEKIIISLADRLLLIHKSLKDAKEFTFDLPGR 529  
DB 296 HNLHRSDETSLSAMKSHA PSGRDGYFLTLRHGNATVDTRVESNMGECTFNVLTPGR 355  
  
QY 530 KYMATVTSISGDLKNSSVKGR----- 552  
DB 356 RYSITVATRSQNLSTSVSEGRGTCPPGSGACPPQCARTVIKSVCCVCVCACVCVCA 415  
  
QY 553 -----VPAQVTDLHVANQMTSSLFTNTWQAQGDVEFYOVLLIHE--- 592  
DB 416 CACVCVCVCAACVRAAVPLERSLTLSSSLG-LGLRASWEKPPGPDVDSVSLTLQDRC 474  
  
QY 593 -----NVVINKESISSSTSYSFHSLKSGSLYSVVVTVTSVSGISSR 633  
DB 475 RRRPVPVQGDGVCVPPSVVQNHSLPVAASSILLSGLTPGALYRLQASTVSGSLRSS 534  
  
QY 634 QVVVEGRTVPSVSGVTNNGRNDYLSVWLVA PGDGVNTEVTLSHDGKVVQSLVIAK- 692  
DB 535 AVSHSGHTSPAASVSDVTVSNGRSDAVRVRPAGVGVVDSYLVRLQDRGRSIHMLAVSR 594  
  
QY 693 SVRSCFSSLTPLGRYTVITTRSKYENHFSQERTVPDKVQGVSVSNARSADYLAVSW 752  
DB 595 SPSECAFSSLVAGRLYTVIVITRSGSLENATVQAQTQPAAVQNPTAVHSARDFLKYVW 654  
  
QY 753 VHATGDFPHYEVTIKNNKFIQTSIKPSKNECVFVQLVPGRLYSVTVTKSGOYEANEQ 812  
DB 655 RHAAGDLDRYLVLRYNLTVLQNSVASQNECVFSSLTPLGRLYTVVETWSGGVSVS 714  
  
QY 813 GNGRTIPEPVKDLTLNRSTEDLHVTWANGVDVQYBIIQLLFNDKMKVPFPFHLVNTATE 872  
DB 715 TDGRTFPAAVRNLSLGDAGTSHLTLTWSAPGDVDHYEVLTFNDTRVFPVVALASQARH 774  
  
QY 873 YRFTSLTPGROYKILVLTISGDVQOASAFIEGFTVPSAVKNHISP-----NGATDSLTVW 928  
DB 775 HRLTSLTPGRHYKMMVSTFSGPYQRAQFLEGRTPSSVKNRLAPLPLGPDVAVGLQASW 834  
  
QY 929 TPGGDDVDSYTVSAFR-HSQKVDSTQTI PKHVFETHRLLEAGEOYQVMIASVSGSLKNQI 987  
DB 835 TPGGDDLDVYVASLSKPDGTVVDRRVPVPKQCSLDFDLVPEGHAYSIAVQSLSGSLTNQS 894

QY 898 NVVGRVTPASVOGVIADNAYSSYSLSVSWOKAAQVABERYDILLTENGILLRNTSEPTAT 1047  
DB 895 AVSRTAPARTALQADNQHTAHSLSVSWEPAGLHAYRLQLLDRGGAVALQRTLPADS 954  
  
QY 1048 KQHFPEDLTPEKKYKIQILTVSGGLFSKEAQTEGRTVPAAVTDLRITENSTRHLSFRWTA 1107  
DB 955 RSEBLEGLTSGRWYRVLTVSGGVAEAEATGQTRPAAVTNLTVASANSSALAFRWOP 1014  
  
QY 1108 SEGELSWNYFLYN-----PDGNLOE----- 1128  
DB 1015 SDGHVDYHVSLYSVPEPAADGRQVSSFRPSPGECGLTEGRGGGRLRVRRPAGGKP 1074  
  
QY 1129 ----RAQVDPLVQSFQNLLOGMYKMWIVTHSG----- 1159  
DB 1075 APPAGGLEPQHE-----QRLHARQNSV---HTGNAALAPPPLVAGGHMTPGVSHAFT 1127  
  
QY 1160 --ELSNESFIFGRVTPASVSHLRGNSRNTTDSLWFMNSPASGDFDYELIYNPNG-TKK 1216  
DB 1128 WRLSSSS-----SSVPSAVSSLEAESQTDGLVVMRHDGDSWGYQVLLCDASGATVA 1183  
  
QY 1217 ENWKDKDLTWRPQGLVPGRYVVLVWVTHSGDLSNKTVAESRTAPSPSLSMSPADIANT- 1275  
DB 1184 VQVVGAAHTSHLFPGLTPGRLYRSEVVTLSGELTNRYSAMGRTAPEPTRLSVRPGATSD 1243  
  
QY 1276 SLAITWKGPPDWTNDYDFELQWLPRDALTVFNPYNRKSEGRIVVGLRPGRSYOFNVKTV 1335  
DB 1244 AIELAWSGPA-LGDYDSFRLWRAPDEPLTVQTH-----LTGRILAGMPPGLNYFTVTV 1298  
  
QY 1336 SGDSWK-----TYSKPIFGSVRTKPKDKIQNLHCRPQNSTAATACSWIPDSDPDGYSIECRK 1391  
DB 1299 SGGASGPGVAISQPIORSVRTSPAPLRSLLHCFPRSSSLSCSWTTPPLSDPDSYEVECR 1358  
  
QY 1392 MDTOEVEFSKRLKBSLLNIMMLVPHKRYLVSIKVSAGMSTSEVEDSTMTIDRPPPP 1451  
DB 1359 QDDGELISVLRLAEGVAATLDHLEAFKYSVTVRLSSAGQTSAPVTHTTVTMIERPVP 1418  
  
QY 1452 PPHIRVNEKDLVLSKSSINFTVNCSEFDTNGAVKYTVVVRADGGDELKPEOHPPLPS 1511  
DB 1419 PPSLOVRSQSKVTSSSILFRNCSWFSDINGAVRFFTVVVAESDADDELLOPEQRHPLPS 1478  
  
QY 1512 YLEVRHNASIRVQNTYFASKCAENPNNSKSFNI-----KLGAEMESIGGKRD 1560  
DB 1479 QDYSRNSLRAVQTAYFPSCCAQSDAPACQOLGINWIIIRQVVEVNLGAGSDRLGGACD 1538  
  
QY 1561 -----PTQOKFCGDLKPHYATYRISIRAPTLQFDEDLKFTKPLXSDTFFSFDIT 1610  
DB 1539 HHRDDDLVLSDSYFPFCGDLKPHYATYRISIRAPTLQFDEDLKFTKPLXSDTFFSFDIT 1598  
  
QY 1611 TE-----SEPLFGAIEGVSAGLFLIGMLVAVVALLICRQVSH- 1648  
DB 1599 THAGMTGTHTRACVHVPHAAVLCABPLGGVVEGSLAQMFLLGTGVVLSLVYRQRLKV 1658  
  
QY 1649 -GRERPSARLSIRDRPLSVHLNLGQGNKRTSCPIKINOFEGHFMKLQADNSYLLSKEY 1707  
DB 1659 AVQENPVVRMSWKEAP-SSGLYVGVSRNRRVSGPVKAGQFOAHLQDLADSSYLLSEEF 1717  
  
QY 1708 EELKDVGRNOSCDIALLPENRGNRYNNILPYDATRVKLSNVDDDDPCSDYINASYI PGN 1767  
DB 1718 EDLKVGRNQAMDVARVPENRGNRYNNILPYDSTRVKLSYLEDDPCSDYINASYVPGNN 1777  
  
QY 1768 FRREYIVTQGPLPOTKDDFWKQWQVQNHVIMVQTCVEKGRVKCDHWPAQDSLYYGD 1827  
DB 1778 -YRREYNPAQGPLPOTKDDFWKQWQVQNHVIMVQTCVEKGRVKCDHWPAQDSLYYGD 1828  
  
QY 1828 LILQMLSESLVPEWTIRREFKICGEQLDAHRLIRHFHYTVMPDRHGVETTSQSLQFVKT 1887  
DB 1829 LVIQKSESZLPEWTIRREFKISSESCGAQPRRLRHFFHYTVMPDRHGVETTSQSLQFVKT 1888  
  
QY 1888 RDIYNRSPGAGTPVHCHSAGVGRGTFTIALDRILQDLSDKSDVDIYGAHVHDLRLHRVHM 1947  
DB 1889 RDIYDRSPSTGATVHCHSAGVGRGTFTIALDRILQDLSDKSDVDIYGAHVHDLRLHRVHM 1948  
  
QY 1948 QTECQYVYLHQCVRDVLRLARKLRSEQENPLFPPIYENVNPEYHR 1990

[illegible]

```
QY 1539 SNSKSFNKLGAEMESLGKGDPTQOKECDGPKLPHHTAYRISIRAFPTQLEDLKEFTKP 1598
Db 121 SSSKSFNKLGAEMDSLGKGDQSQKFCDFGLPLPHHTAYRISIRAFPTQLEDLKEFTKP 180
QY 1599 LYSTDFSLPTTTESEPLFGAIEGVSAFLIGMLVAVALLICRQKVSHGRERPSARLS 1658
Db 181 LYSTDFSMPTTTESEPLFGVIEGVSAFLIGMLVALVAFFICRQKASHRERPSARLS 240
QY 1659 IRRDRPLSVHLNLQKGNKRTSCPIKINQFEGHFMKLOADSNLLSKEYEELKDVGRNQ 1718
Db 241 IRRDRPLSVHLNLQKGNKRTSCPIKINQFEGHFMKLOADSNLLSKEYEELKDVGRSQ 300
QY 1719 CDIALLENKRNKRYNLPYDAFRVLSNVDDPCSDYNASYPGNFRFRIYVQGP 1778
Db 301 CDIALLENKRNKRYNLPYDASRVLSNVDDPCSDYNASYPGNFRFRIYVQGP 360
QY 1779 LPGTGKDFWKMVQNVHNIWMVTQCVKEGKRVKCDHYWPAQDQSLYGLDLIQLMSESVL 1838
Db 361 LPGTGKDFWKMVQNVHNIWMVTQCVKEGKRVKCDHYWPAQDQSLYGLDLIQLMSESVL 420
QY 1839 PEWTIREFKICSEBQDLAHLRIHFHYTVWPDHGVPTTQSLIQFVRTVRYINRSPGAG 1898
Db 421 PEWTIREFKICSEBQDLAHLRIHFHYTVWPDHGVPTTQSLIQFVRTVRYINRSPGAG 480
QY 1899 PTVVHCSAGVGTGTFALDRILQOLSKDSVDIYGAVHDLRLHRVHMVQTECOYVYLHQ 1958
Db 481 PTVVHCSAGVGTGTFALDRILQOLSKDSVDIYGAVHDLRLHRVHMVQTECOYVYLHQ 540
QY 1959 CVRDVLRAKLRSQENPLPIYENVNPEYHRDPVYSRH 1997
Db 541 CVRDVLRAKLRSQENPLPIYENVNPEYHRDAIYSRH 579

RESULT 10
Q86VA4 HUMAN
ID Q86VA4 HUMAN PRELIMINARY; PRT; 771 AA.
AC Q86VA4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PTPRB protein.
GN Name=PTPRB;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
TI TISSUE=Spleen;
ME MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwain P.J., McKernan K.J., Malek J., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
```

```
RP NUCLEOTIDE SEQUENCE.
TI TISSUE=Spleen;
RA Director MGC Project;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC051329; AAH51329.1; -; mRNA.
DR Ensembl; ENSG00000127329; Homo sapiens.
DR GO; GO:0005529; F-sugar binding; IEA.
DR InterPro; IPR003961; FN.III.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00041; fn3; 6.
DR Pfam; PF00652; Ricin B lectin; 1.
DR SMART; SM00060; FN3; 6.
DR PROSITE; PS00853; FN3; 6.
DR PROSITE; PS02031; RICIN_B_LECTIN; 1.
SQ SEQUENCE 771 AA; 86055 MW; B6E388D6E3F573BF CRC64;

Query Match 26.4%; Score 2772; DB 2; Length 771;
Best Local Similarity 99.4%; Pred. No. 8.1e-153;
Matches 533; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 18 QTGLAEPERCNFTLAESKASHSVSIQWRILGSPCNFSLIYSSDTLGAALCPTFRIDNTT 77
Db 236 ETGLAEPERCNFTLAESKASHSVSIQWRILGSPCNFSLIYSSDTLGAALCPTFRIDNTT 295
QY 78 YGCNLDQLQAGTIYNFKIISLDEERTVVLQDPLPPARFGVSKETTTSTGLHVMWTPSSG 137
Db 296 YGCNLDQLQAGTIYNFKIISLDEERTVVLQDPLPPARFGVSKETTTSTLHVMWTPSSG 355
QY 138 KVTSEYVQLFDENNKKIQGVQIQBSTSWNEVTFENLTAGSKYNIATAVSGGKESFSVYT 197
Db 356 KVTSEYVQLFDENNKKIQGVQIQBSTSWNEVTFENLTAGSKYNIATAVSGGKESFSVYT 415
QY 198 NGSTVPSPVKDIGISTKANSLISWSHGSGNVERYLMLMDKGLVHGGVVDKHAITSYAP 257
Db 416 NGSTVPSPVKDIGISTKANSLISWSHGSGNVERYLMLMDKGLVHGGVVDKHAITSYAP 475
QY 258 HGLSPGYLYNLTVMTAEAGLQNYRWKLVRTAPMEVSNLKVTDGSLTSKVKWQRPQGNV 317
Db 476 HGLTPGYLYNLTVMTAEAGLQNYRWKLVRTAPMEVSNLKVTDGSLTSKVKWQRPQGNV 535
QY 318 DSYNITLSHGKTIKESRVLPWITETHPKELVLPQRLYQVTVSCVSGELSAQKMAVGRTFP 377
Db 536 DSYNITLSHGKTIKESRVLPWITETHPKELVLPQRLYQVTVSCVSGELSAQKMAVGRTFP 595
QY 378 DKVANLEANNNGRMRSLLVSVSPAGDWEQYRILLFNDSVVLNITVYKKEETQYVMDDTG 437
Db 596 DKVANLEANNNGRMRSLLVSVSPAGDWEQYRILLFNDSVVLNITVYKKEETQYVMDDTG 655
QY 438 LVPGRQYEVVIVESGNLKNSERCQGRTPVLAVLQLRVKHANETSLSIMWQTPVAWEKY 497
Db 656 LVPGRQYEVVIVESGNLKNSERCQGRTPVLAVLQLRVKHANETSLSIMWQTPVAWEKY 715
QY 498 IISLADRLLLIHKLSKDAKEFTFDLPGRKYMATVTSISGLKNSSSVKGRTV 553
Db 716 IISLADRLLLIHKLSKDAKEFTFDLPGRKYMATVTSISGLKNSSSVKGRTV 771

RESULT 11
Q6ZTX7 HUMAN
ID Q6ZTX7 HUMAN PRELIMINARY; PRT; 609 AA.
AC Q6ZTX7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE CDNA FLJ44133 f1s, clone THYMU2008725, highly similar to PROTEIN-
TYROSINE PHOSPHATASE BETA (EC 3.1.3.48).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
```

```
RC TISSUE=Thymus;
RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA Arita M., Mueashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA Ono T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
RA Isoku Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Murakawa K., Kanehori K., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isogai T.,
RT "NEDO human cDNA sequencing project.";
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK126121; BAC86450.1; -; mRNA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR00772; Ricin_B_lectin.
DR Pfam; PF00041; fn3; 4.
DR Pfam; PF00652; Ricin_B_lectin; 1.
DR SMART; SM00060; FN3; 4.
DR PROSITE; PS00853; FN3; 4.
DR PROSITE; PS0231; RICIN_B_LECTIN; 1.
SQ SEQUENCE 609 AA; 67816 MW; 268E3682457E0660 CRC64;

Query Match 17.7%; Score 1854.5; DB 2; Length 609;
Best Local Similarity 93.0%; Pred. No. 1.9e-99;
Matches 356; Conservative 5; Mismatches 3; Indels 19; Gaps 1;

Oy 18 QTGLAEPERCNFTLAESKASHSVSIQWRILGSPCNFSLIYSSDTLGAALCPTFRIDNTT 77
Db 236 ETGLAEPERCNFTLAESKASHSVSIQWRILGSPCNFSLIYSSDTLGAALCPTFRIDNTT 295

Oy 78 YGCMLODQAGTYNFKIISLDEBRTVVLQDPLPPARFGVSKETKSTGLHVVWTPSSG 137
Db 296 YGCMLODQAGTYNFKIISLDEBRTVVLQDPLPPARFGVSKETKSTGLHVVWTPSSG 355

Oy 138 KVTSEYQLPDENNQKIQGVQIOESTSWNEYTRFNLTAGSKYNIAITAVSGKRSFVYT 197
Db 356 KVTSEYQLPDENNQKIQGVQIOESTSWNEYTRFNLTAGSKYNIAITAVSGKRSFVYT 415

Oy 198 NGSTVPSPVKDIGISTKANSLLISWSHSGNVERYRLMDKGLVHGGVVDKHTSYAF 257
Db 416 NGSTVPSPVKDIGISTKANSLLISWSHSGNVERYRLMDKGLVHGGVVDKHTSYAF 475

Oy 258 HGLSPGVLNLTWTEAAGLQNTKWLVRTPAPMEVSNLKVTDGSLTSLKVKQRPNGV 317
Db 476 HGLTPGVLNLTWTEAAGLQNTKWLVRTPAPMEVSNLKVTDGSLTSLKVKQRPNGV 535

Oy 318 DSYNTILSHKGTIKESRLVAPWITETHFKELVPCRLYQVTVSCVSGELSAQKMAVGRTFP 377
Db 536 DSYNTILSHKGTIKESRLVAPWITETHFKELVPCRLYQVTVSCVSGELSAQKMAVGRT-- 593

Oy 378 DKVANLEANNGRMRSILVWSWSP 400
Db 594 -----CESWAP 599

RESULT 12
O9CX77 MOUSE
ID O9CX77 MOUSE PRELIMINARY; PRT; 723 AA.
AC O9CX77;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length enriched
DE library, clone:3230402H02 product:hypothetical fibronectin type III
DE domain/Ricin B lectin domain containing protein, full insert
DE sequence.
GN Name=Ptpzb;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1;
RP NUCLEOTIDE SEQUENCE.
```

```
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinegawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gisi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schraml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Bargh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Maehima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head;
RA The PANTOM Consortium.
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki H., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
```

```
RA Muramatsu M., Hayaehizaki Y.;
PL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK019450; BAB1176.1; -, mRNA.
DR Ensembl; ENSMUSG00000020154; Mus musculus.
DR MGI; MGI:97809; Ptpnb.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR000772; R1cin_B_lectin.
DR Pfam; PF00041; fn3; 5.
DR SMART; SM00060; FN3; 5.
DR PROSITE; PS00853; FN3; 5.
DR PROSITE; PS0231; RICIN B LECTIN; 1.
KW Hypothetical protein: Lectin.
SQ SEQUENCE 723 AA; 79341 MW; 184F7E96B32EEF CRC64;

Query Match 16.4%; Score 1718.5; DB 2; Length 723;
Best Local Similarity 72.6%; Pred. No. 2.2e-91;
Matches 326; Conservative 50; Mismatches 72; Indels 1; Gaps 1;

Qy 18 QTGLAEPERCNFTLAESKASHSVSIQWRLGSCNFSLIYSSDTLGAALCPTFRIDNTT 77
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
266 ETGVAEQKCNFTLLESRSVLSASIQWRTFASPNCNFSLIYSSDTSGPMWCHPIRIDNFT 325
Qy 78 YGCNLQDLQAGTIYNEKILSID-BERTVLTQDPLPPARFGVSKETTTSTGLHVWTPSS 136
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
326 YGCNPKDLQAGTVNFRIVSLDGEESTLVLTQDPLPPARFEVNSEKTASTTLQVRWTPSS 385
Qy 137 GKVTSEYVQLFDENNQIKQVIOESTSWNEYTEFNLTAGSKYNIATVAVSGGKRSPSVY 196
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
386 GKVSWEYVQLFDHNDKIQEVOVESTWSTQYTLNITEGNSYKVAITAVSGEKRSPVY 445
Qy 197 TNGSTVSPVKDIGISTKANSLISWSHGSGNRYRLMLDKGILVHGGVVDKHAATSYA 256
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
446 INGSTVSPVKDIGISPNNSLLISWSGSGNRYRLMLDKGILVHGGVVDKHAATSYA 505
Qy 257 FHGLSPGYLNTVMTAEAGLQNYRWKLVRTAPNEVSNLKVNDGSLTSLKVKWQRPNGN 316
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
506 FHGLTPGHLNLTVTWASGLNSRWKLVRTAPNEVSNLKVNDGRLTSLNVRWQKPGPD 565
Qy 317 VDSNYITLSHGTTIKESRGLAPMTTETHFKELVPGRLYQVTVSCVSGELSAQKMAVGRTF 376
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
566 VDSNYITLSHGTTIKESRGLAPMTTETHFKELVPGRLYQVTVSCVSGELSAQKMAVGRTV 625
Qy 377 PDKVANLEANNGMRSLVSWSPAGDWEQYRILLFNDSVLLNITVGKEETQYVMDT 436
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
626 PEKVRNLVSYNEIWMKSFTVNTWTPAGDWEHYRIVLNFSLVLLNTTVGKEETHYALDGL 685
Qy 437 GLVPGROYEVEVIVESGNLKNRSCQGR 465
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
686 ELIPGROYEVEVIVESGNLKNRSCQGR 714

RESULT 13
Q7PQLO ANOGA PRELIMINARY; PRT; 1506 AA.
AC Q7PQLO;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000013354 (Fragment).
GN ORFNames=ENSANGG0000010865;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

[2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008880; EAA08669.3; -, Genomic DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004896; F:hematopoietin/interferon-class (D200-domain. .; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.
DR InterPro; IPR000282; Cytok_receptor_2.
DR InterPro; IPR003962; FNIII_subd.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR003959; PTPC motif.
DR InterPro; IPR000387; TYR_phosphatase.
DR InterPro; IPR000242; Tyr_PP.
DR Pfam; PF00041; fn3; 10.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR PRINTS; PR00700; PRTYPHPTASE.
DR SMART; SM00060; FN3; 11.
DR SMART; SM00194; PTPC; 1.
DR SMART; SM00404; PTPC_motif; 1.
DR PROSITE; PS00853; FN3; 11.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00055; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase.
FT NON TER 1 1
FT NON TER 1506 1506
SQ SEQUENCE 1506 AA; 171796 MW; CF92A08B60EEA5DE CRC64;

Query Match 14.9%; Score 1566; DB 2; Length 1506;
Best Local Similarity 29.5%; Pred. No. 6e-82;
Matches 512; Conservative 263; Mismatches 583; Indels 376; Gaps 74;

Qy 342 ETHFKELVPGRLYQV-----TVSCVSGELSAQKMAVGRTFPDKVANLEA- 385
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
43 EIQFNGCLPQTRYNFWLYTNTSTHKLWLTWVSIT-----TAPDPPANLTVI 89
Qy 386 NNNGMRSLVSWSPPA-GWEOY--RILLFNDSVLLNITVGKEETQYVMDTGLVPGR 442
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
90 PRSG--KNVIINSPPAQGNYSFSLKILGLSDNFA-TNQTVAIEDNQFQYVMDLTPGA 146
Qy 443 QYEVE-----VIVESGNLKNRSCQGRTPVLAQLRVKHANETSLSIMWOT--P 490
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
147 TYQVQAVTLYDGKESVAYTSRNFTTKPTPKGFI-----VWFRNETLLVLMQPPYP 198
Qy 491 VAWEKVIISLADRLILLIKSLSKDAK-----EFTFDLVPGRKYNMATVTSISGD-LKN 544
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
199 AGIYTHYKVSIEPPDALGSLVYVQKEGEPGPAQAAGKGLVPGRAYNISVQTHSEDSL 258
Qy 545 SSSVKGRTVPAQVTDLHVANQGMTSSLF-TNWTQAQGDVEF--YQVLLI---HENVVIKN 598
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
259 PTTAQYRTVPLRPMNVTFDKKSITENSFKVWEAPKGTSEFQYOVSLSTRRQQAVALRN 318
Qy 599 ESISSETSRYSFHSLSGSLYVVVVTVVSGGISRQVQVVEGRTVPSSVSGVTVNNSGRND 658
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
319 DNENMAWLEFK-DNLDPGTKYQVWTVSGKV----- 350
Qy 659 YLSVSMIVAPGDVDNVEVTLSDGKVVQSLVIAKSVRECSFSLTPGLRTVTVITTSRSGK 718
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
351 ----SW-PASGD-----VTL-----KPLPVKQLQSYTDSKTGVILLISWK--- 384
Qy 719 YENHSFQERTVPDKVQGSVNSARSADYLRLSVNWHATGPDHYEVTIKNNPIQTKSI 778
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
385 -----PDEL-----STQDEY-RIS-----YHELETNNGD-----SS 409
Qy 779 PKGENECFV--QLVPGRLYSVTVTTKSGQYEANE-----QNGRTIPEPVKDLTLNR 830
```



RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RA "The genome sequence of *Drosophila melanogaster*";  
RL Science 287:2185-2195(2000).  
RN [4]  
RP GENOME REANNOTATION, AND ALTERNATIVE SPLICING.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminler J.S., Millburn G.H., Prochuk S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
RT systematic review";  
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
RN [5]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM LONG).  
RC STRAIN=Berkeley; TISSUE=Embryo;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.W.,  
RA Champe M., Chavez C., Dorsett V., Dresner D., Farfan D., Frise E.,  
RA George R.A., Gonzalez M., Guarin H., Kronmiller B., Li P.W., Liao G.,  
RA Miranda A., Mungall C.J., Nuncio J., Paclet J.M., Paragas V., Park S.,  
RA Patel S., Phouanavong S., Wan K.H., Yu C., Lewis S.E., Rubin G.M.,  
RA Celniker S.E.;  
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein  
CC tyrosine + phosphate.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=4;  
CC Name=Long;  
CC IsoId=P35992-1; Sequence=Displayed;  
CC Name=Short;  
CC IsoId=P35992-2; Sequence=VSP\_005143, VSP\_005144;  
CC Name=A; Synonyms=B;  
CC IsoId=P35992-3; Sequence=VSP\_015266;  
CC Note=No experimental confirmation available;  
CC Name=C;  
CC IsoId=P35992-4; Sequence=VSP\_015264, VSP\_015265;  
CC Note=No experimental confirmation available;  
CC -1- TISSUE SPECIFICITY: Selectively expressed in anterior commissure  
CC and its junctions with the longitudinal tracts.  
CC -1- SIMILARITY: Belongs to the protein-tyrosine-phosphatase family.  
CC Receptor class subfamily  
CC -1- SIMILARITY: Contains 12 fibronectin type-III domains.  
CC -1- SIMILARITY: Contains 12 fibronectin type-III domains.  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC EMBL; M80465; AAA28484.1; -; mRNA.  
CC EMBL; M80538; AAA28952.1; -; mRNA.  
CC EMBL; AE003486; AAF48072.2; -; Genomic DNA.  
CC EMBL; AE003486; AAO09638.1; -; Genomic DNA.  
CC EMBL; BT004474; AAO42638.1; -; mRNA.  
CC PIR; D41214; D41214.  
CC HSP; P10586; ILAR.  
CC Ensembl; CG1817; *Drosophila melanogaster*.  
CC Flybase; FBgn004370; Ftp10D.  
GO; GO:0004725; F:protein tyrosine phosphatase activity; IDA.

DR GO; GO:0006470; P:protein amino acid dephosphorylation; IDA.  
DR InterPro; IPR000282; Cytok\_receptor\_2.  
DR InterPro; IPR003961; FN\_III.  
DR InterPro; IPR003962; FNIII subd.  
DR InterPro; IPR000387; TYR\_phosphatase.  
DR InterPro; IPR000242; Tyr\_PP.  
DR Pfam; PF00041; fn3\_10.  
DR Pfam; PF00102; Y\_phosphatase; 1.  
DR PRINTS; PR00014; FNTYPEIII.  
DR PRINTS; PR00700; PRTYPHPTASE.  
DR SMART; SM00060; FN3; 11.  
DR SMART; SM00194; PTPc; 1.  
DR PROSITE; PS00853; FN3\_11.  
DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
DR PROSITE; PS00056; TYR\_PHOSPHATASE\_2; 1.  
DR PROSITE; PS00055; TYR\_PHOSPHATASE\_PTP; 1.  
KW Alternative splicing; Glycoprotein; Hydrolase; Protein phosphatase;  
KW Repeat; Signal; Transmembrane.  
FT SIGNAL 1 42 Potential.  
FT CHAIN 43 1631 Tyrosine-protein phosphatase 10D.  
FT TOPO\_DOM 43 1197 Extracellular (Potential).  
FT TRANSMEM 1198 1218 Potential.  
FT TOPO\_DOM 1219 1631 Cytoplasmic (Potential).  
FT DOMAIN 43 119 Fibronectin type-III 1.  
FT DOMAIN 120 214 Fibronectin type-III 2.  
FT DOMAIN 215 308 Fibronectin type-III 3.  
FT DOMAIN 309 402 Fibronectin type-III 4.  
FT DOMAIN 403 493 Fibronectin type-III 5.  
FT DOMAIN 494 580 Fibronectin type-III 6.  
FT DOMAIN 581 669 Fibronectin type-III 7.  
FT DOMAIN 670 766 Fibronectin type-III 8.  
FT DOMAIN 767 861 Fibronectin type-III 9.  
FT DOMAIN 862 955 Fibronectin type-III 10.  
FT DOMAIN 956 1048 Fibronectin type-III 11.  
FT DOMAIN 1049 1197 Fibronectin type-III 12.  
FT DOMAIN 1272 1520 Tyrosine-protein phosphatase.  
FT ACT\_SITE 1468 1468 Phosphocysteine intermediate (By  
FT similarity).  
FT CARBOHYD 75 75 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 106 106 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 128 128 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 169 169 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 212 212 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 229 229 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 259 259 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 289 289 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 317 317 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 471 471 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 486 486 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 512 512 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 533 533 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 588 588 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 668 668 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 687 687 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 719 719 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 723 723 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 823 823 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 841 841 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 874 874 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 908 908 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 925 925 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 1001 1001 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 1104 1104 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 1136 1136 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 1195 1195 N-linked (GlcNAc... ) (Potential).  
FT VARSPIC 1549 1558 GQQVLDENG -> DDEGIAESGM (in isoform  
FT Short).

Query Match 14.9%; Score 1561.5; DB 1; Length 1631;  
Best Local Similarity 27.9%; Pred. No. 1.3e-81;  
Matches 514; Conservative 271; Mismatches 633; Indels 421; Gaps 74;  
Cy 265 LYNLTVMTEAAGLQ-----NYRW-----KLVRTAPMEVS--NLKVTNDGSLT 304







Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 April J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
 Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 Borkova D., Botchan M.R., Bouch J., Brokstein P., Brottier P.,  
 Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 Cherry J.K., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 Harris N.L., Harvey D.A., Helman T.J., Hernandez J.R., Houck J.,  
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
 Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 Merkulov G., Milahina N.V., Mobarry C., Morris J., Moshrefi A.,  
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 "finishing a whole-genome shotgun: release 3 of the Drosophila  
 melanogaster euchromatic genome sequence."  
 Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 [2]  
 NUCLEOTIDE SEQUENCE.  
 MEDLINE=22426065; PubMed=12537568;  
 Celinker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
 Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
 "finishing a whole-genome shotgun: release 3 of the Drosophila  
 melanogaster euchromatic genome sequence."  
 Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 [3]  
 NUCLEOTIDE SEQUENCE.  
 MEDLINE=22426070; PubMed=12537573;  
 Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,  
 Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 Ashburner M., Celinker S.E.,  
 "the transposable elements of the Drosophila melanogaster euchromatin:  
 a genomic perspective."  
 Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).  
 [4]  
 NUCLEOTIDE SEQUENCE.  
 MEDLINE=22426069; PubMed=12537572;  
 Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 Bettencourt B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A.,  
 Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 Lewis S.E.,  
 "Annotation of the Drosophila melanogaster euchromatic genome: a  
 systematic review."  
 Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 [5]  
 NUCLEOTIDE SEQUENCE.  
 Berkeley Drosophila Genome Project;  
 Celinker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,  
 Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,  
 Yu C., Rubin G.

"Drosophila melanogaster release 4 sequence."  
 Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 [6]  
 RT NUCLEOTIDE SEQUENCE.  
 RN Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
 RG -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC EMBL; AE003432; AN09133.1; -, Genomic\_DNA.  
 DR HSP86; P10586; IJAR.  
 DR Ensembl; CG8899; Drosophila melanogaster.  
 DR FlyBase; FBgn004368; Ctp4E.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0004725; F:protein tyrosine phosphatase activity; IEA.  
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.  
 DR InterPro; IPR003962; FNIII subd.  
 DR InterPro; IPR003961; FN III.  
 DR InterPro; IPR000387; TYR phosphatase.  
 DR InterPro; IPR000242; Tyr\_PP.  
 DR Pfam; PF00041; fn3; 10.  
 DR Pfam; PF00102; Y phosphatase; 1.  
 DR PRINTS; PR00014; FNTYPEIIII.  
 DR PRINTS; PR00700; PRTYDPHTASE.  
 DR SMART; SM00060; FN3; 11.  
 DR SMART; SM00394; PTEC; 1.  
 DR PROSITE; PSS0853; FN3; 11.  
 DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; 1.  
 DR PROSITE; PS00056; TYR\_PHOSPHATASE\_2; 1.  
 DR PROSITE; PSS0055; TYR\_PHOSPHATASE\_ptp; 1.  
 KW Hydrolase; Repeat; Transmembrane.  
 SQ SEQUENCE 1607 AA; 182285 MW; 6CAF554E61A27F45 CRC64;  
 Query Match 14.2%; Score 1485.5; DB 2; Length 1607;  
 Best Local Similarity 29.2%; Pred. No. 3.3e-77;  
 Matches 475; Conservative 269; Mismatches 687; Indels 197; Gaps 62;  
 Qy 456 KNSERCGRVPLAVLQLRV---KHANETSLINWQTPVAWEK---YIISLA-----D 503  
 Db 44 QNSQQQQRHFVWVVGILITFLAQHANADLV--NVPNASSNANAFYRIDYSPFGPPE 101  
 Qy 504 RDLIIHKLSKDAKEFTFDLPGRKY--MATVTISGDLKNSSSVKGRTPVAQVTDLH 561  
 Db 102 PNTTIPASDIGNKDIK---FSRALPGTYNFWLYTNTSHREQLTWTNITAPDPANLS 158  
 Qy 562 VANQGMTSSLFTNWTOAQGVDFYVQLLIHENVVKNESIS---SETSRYSFHSKSGSL 618  
 Db 159 VQLRSSKSAITWRPPGSGRYSGFRIRVLGTLDPERSYSLEGNETLQSAKELTPGS 218  
 Qy 619 YSVVTVTVSGGISRRQVVEGRTPVPSVSGVTNNNSGRNDYLSVSWL--VAPGVDNVEV 676  
 Db 219 YQVQAVSVYQKESVAYTSRNFTHKNTPGKFIWFRNETTLLVLWQPPFAGIYTHRV 278  
 Qy 677 TLSHDGKVVQSLVIK-----SVRECSFSLTCLRLYTVTTTTSKGVNHSFS-----Q 726  
 Db 279 SITPDDAIQSVLYVEREGEPGPAQAAKGLVGRBNISVQTVS---EDETSSVPTTAR 335  
 Qy 727 ERTVPDKVQGVSVNS--ARSDYLRVSW--VHATGDFHYEVTIKNKNFIQTSKIPKSEN 783  
 Db 336 YLTVPERVLNVTDEAVTTSSFRVWEPPRTYSEFDAYQVWLSTSRIP---NVPRAAN 392  
 Qy 784 -ECVFTVQ-----LVPGRLYSVTVTTKSGQYBANEGNORTI-PEPVKOL--TLNRSTEDL 835  
 Db 393 GDSVYFDYDILPGRTEYVWVTKTIADNVNNSWSPASGEVTLRPRVRSGLGFLDNRSA-L 451  
 Qy 836 HVTWMSGA-NGVDVQYEIQL--LFNDMKVFPPE-----HLVNTATYERFTSLTPGROYKI 886  
 Db 452 HISWEPAETGRQDSYRISVHEQTNASVPAPFVPAESQITNTLTLSLQSLAGRIYLI 511  
 Qy 887 LVLITISGDVQQSAP-IEGFTVPSPA--VKNTHISPNGATDSLTVNWTGGGDV-----DSYT 939  
 Db 512 AVQALSKGVASNADITRYTPAPLIQELRSIDQG-----LMLSWR---SDVNSRQDRYE 564



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2006, 18:36:34 ; Search time 49.7221 Seconds  
(without alignments)  
3320.525 Million cell updates/sec

Title: US-10-633-742-2

Perfect score: 10483

Sequence: 1 MLSHGAGLALWITLSLLQTG.....PPIYENVNPEVHRDPVYSRH 1997

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5 COMB pep:\*
- 2: /cgn2\_6/ptodata/1/iaa/6 COMB pep:\*
- 3: /cgn2\_6/ptodata/1/iaa/H COMB pep:\*
- 4: /cgn2\_6/ptodata/1/iaa/PTCUS COMB pep:\*
- 5: /cgn2\_6/ptodata/1/iaa/RE COMB pep:\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10483	100.0	1997	2	US-09-949-016-6275
2	9898	94.4	1903	2	US-09-949-016-8049
3	1444	13.8	278	1	US-08-201-697-16
4	1374	13.1	254	1	US-08-685-992-14
5	1374	13.1	254	1	US-09-144-925-14
6	1247	11.9	1337	2	US-08-854-585-2
7	1247	11.9	1337	2	US-09-447-533-2
8	1247	11.9	1337	4	PCT-US95-05512-2
9	1198	11.4	1711	1	US-08-342-930-2
10	1107	10.6	2291	1	US-08-822-871-2
11	1097	10.5	2301	1	US-08-822-871-4
12	915.5	8.7	1188	2	US-08-201-697-4
13	913.5	8.7	1246	2	US-09-949-016-8051
14	913.5	8.7	1246	2	US-09-949-016-8052
15	913.5	8.7	1246	2	US-09-949-016-8053
16	913.5	8.7	1246	2	US-09-949-016-8054
17	913.5	8.7	1246	2	US-09-949-016-8055
18	908.5	8.7	1187	1	US-08-201-697-2
19	899.5	8.6	1274	2	US-09-949-016-8828
20	899.5	8.6	1274	2	US-09-949-016-8829
21	899.5	8.6	1274	2	US-09-949-016-8830
22	899.5	8.6	1274	2	US-09-949-016-8831
23	899.5	8.6	1274	2	US-09-949-016-8832
24	843.5	8.0	273	1	US-08-201-697-17
25	817	7.8	898	1	US-08-036-210-22
26	817	7.8	898	1	US-08-449-609-22
27	817	7.8	898	2	US-08-361-096A-22

28	796	7.6	506	2	US-09-949-016-8833	Sequence 8833, Ap
29	796	7.6	506	2	US-09-949-016-8834	Sequence 8834, Ap
30	796	7.6	506	2	US-09-949-016-8835	Sequence 8835, Ap
31	796	7.6	506	2	US-09-949-016-8836	Sequence 8836, Ap
32	796	7.6	506	2	US-09-949-016-8837	Sequence 8837, Ap
33	796	7.6	506	2	US-09-949-016-8838	Sequence 8838, Ap
34	761.5	7.3	1911	1	US-08-348-006B-5	Sequence 5, Appli
35	761.5	7.3	1911	1	US-08-800-825A-5	Sequence 5, Appli
36	761.5	7.3	1911	2	US-09-158-857-5	Sequence 5, Appli
37	761.5	7.3	1911	4	PCT-US94-10166-5	Sequence 3, Appli
38	748.5	7.1	1501	1	US-08-447-464-3	Sequence 3, Appli
39	748.5	7.1	1501	1	US-08-716-679-3	Sequence 3, Appli
40	744.5	7.1	251	1	US-08-685-992-15	Sequence 15, Appli
41	744.5	7.1	251	1	US-09-144-925-15	Sequence 15, Appli
42	719.5	6.9	1452	1	US-08-652-371-4	Sequence 4, Appli
43	719.5	6.9	1452	1	US-08-991-258A-4	Sequence 4, Appli
44	719.5	6.9	1452	1	US-08-769-399-4	Sequence 4, Appli
45	719.5	6.9	1452	2	US-08-991-953A-4	Sequence 4, Appli

#### ALIGNMENTS

#### RESULT 1

US-09-949-016-6275  
; Sequence 6275, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6275  
; LENGTH: 1997  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-6275

Query Match		100.0%;	Score 10483;	DB 2;	Length 1997;
Best Local Similarity		100.0%;	Pred. NO. 0;		
Matches 1997;		Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Qy	1	MLSHGAGLALWITLSLLQTGLABPERCNFTLAESKASHSVSIQWRILGSPCNFSLIYSS	60		
Db	1	MLSHGAGLALWITLSLLQTGLABPERCNFTLAESKASHSVSIQWRILGSPCNFSLIYSS	60		
Qy	61	DTLGAALCPTFRIDNTYTCNQLQDQAGTIYFNKIIISLDEERTVVLQDPLPPARFGVSK	120		
Db	61	DTLGAALCPTFRIDNTYTCNQLQDQAGTIYFNKIIISLDEERTVVLQDPLPPARFGVSK	120		
Qy	121	EKTSTGLHWVWPSSGKVTSYEVQLFDENNQIQGVQIQUESTSWNYYTFNLTAGSKYN	180		
Db	121	EKTSTGLHWVWPSSGKVTSYEVQLFDENNQIQGVQIQUESTSWNYYTFNLTAGSKYN	180		
Qy	181	IATFVSGGKRSPVYNTGSTVPSPVKDIGISTKANSLLSWSHSGSNVRYRLMLMDKG	240		
Db	181	IATFVSGGKRSPVYNTGSTVPSPVKDIGISTKANSLLSWSHSGSNVRYRLMLMDKG	240		
Qy	241	ILVHGGVVDKHTASYAFHGLSPGYLYNLTWMTAAGLQNYRWKLVRTAPMEVSNLKVTD	300		
Db	241	ILVHGGVVDKHTASYAFHGLSPGYLYNLTWMTAAGLQNYRWKLVRTAPMEVSNLKVTD	300		
Qy	301	GSLSLSLKVKWQRPNGVDSYNTLSLHGKTIKESRVLAPWITETHFKEVLPGRLYQVTVSC	360		

Db 301 GSLTSLVKWQPGNVDNYITLSHGKTIIESRVLAPWITETHFKELVPGRLQVTVSC 360  
Qy 361 VSGELSAQMAVGRTFPDKVANLANNGRMRSLSLVWSMSPAGDWEQVRIILLFNDSSVLL 420  
Db 361 VSGELSAQMAVGRTFPDKVANLANNGRMRSLSLVWSMSPAGDWEQVRIILLFNDSSVLL 420  
Qy 421 NITVKEBTQVMDDTGLVPGRQYEVIVBSGNLKNSERCQGRTPVLAVLQLRVGHANE 480  
Db 421 NITVKEBTQVMDDTGLVPGRQYEVIVBSGNLKNSERCQGRTPVLAVLQLRVGHANE 480  
Qy 481 TSLSIMQTPVAWEKYYIISLADRDLLIHKSLSKDAKEFTPTDLVPGRKYMATVTSISG 540  
Db 481 TSLSIMQTPVAWEKYYIISLADRDLLIHKSLSKDAKEFTPTDLVPGRKYMATVTSISG 540  
Qy 541 DLKSSSVKGRTPVAQVTDLHVANQGMTSSLFTWNTQAQGDVEFYQVLLIHENVVIKNES 600  
Db 541 DLKSSSVKGRTPVAQVTDLHVANQGMTSSLFTWNTQAQGDVEFYQVLLIHENVVIKNES 600  
Qy 601 ISSSETSRYSFHSLSKSGSLYSVVVTTVSGGISRRQVWVEGRTVPSSVSGVTVNNSGRNDYL 660  
Db 601 ISSSETSRYSFHSLSKSGSLYSVVVTTVSGGISRRQVWVEGRTVPSSVSGVTVNNSGRNDYL 660  
Qy 661 SVSMLVAPGDVDNTEVTLSDHDKVQVSLVIAKSVRECSFSLTPGRLYTVTITTRSGKYE 720  
Db 661 SVSMLVAPGDVDNTEVTLSDHDKVQVSLVIAKSVRECSFSLTPGRLYTVTITTRSGKYE 720  
Qy 721 NHSFSQERTVPDKVQVSVNSARSDDLVRVSWHATGDFHVEYTIKNKNFIQTKSIK 780  
Db 721 NHSFSQERTVPDKVQVSVNSARSDDLVRVSWHATGDFHVEYTIKNKNFIQTKSIK 780  
Qy 781 SENECSFVQLVPGRLYSVTVTTKSGQYEAENGNGRTIPEPKDILTNRSTEDLHVTVS 840  
Db 781 SENECSFVQLVPGRLYSVTVTTKSGQYEAENGNGRTIPEPKDILTNRSTEDLHVTVS 840  
Qy 841 GANGDVQYEQIQLLFNDKMFVPPPHLVNTATEYFTSLTPGRQYKILVLITSGDVQOSAF 900  
Db 841 GANGDVQYEQIQLLFNDKMFVPPPHLVNTATEYFTSLTPGRQYKILVLITSGDVQOSAF 900  
Qy 901 IEGFTVPSAVKNIIHISNGATDSLTVNMTFGGVDVSTVSAPRHSQKVDQSITIPKHVE 960  
Db 901 IEGFTVPSAVKNIIHISNGATDSLTVNMTFGGVDVSTVSAPRHSQKVDQSITIPKHVE 960  
Qy 961 HTFHLRLEAGEQYQIMIASVSGSLKNQINNVGRTVPASVQGVIAADNAYSSYSLIVSMQKAA 1020  
Db 961 HTFHLRLEAGEQYQIMIASVSGSLKNQINNVGRTVPASVQGVIAADNAYSSYSLIVSMQKAA 1020  
Qy 1021 GVAERYDILLTENGILLRNTSEPAATTKQHKFEDLTPGKYYKIQLTVSGGLFSKEAQTE 1080  
Db 1021 GVAERYDILLTENGILLRNTSEPAATTKQHKFEDLTPGKYYKIQLTVSGGLFSKEAQTE 1080  
Qy 1081 GRTVPAAVTDLRIITENSTRHLSFWMTASEGELSWYNIFLYNDPQNLQERAOVDPLVQSPS 1140  
Db 1081 GRTVPAAVTDLRIITENSTRHLSFWMTASEGELSWYNIFLYNDPQNLQERAOVDPLVQSPS 1140  
Qy 1141 FQNLLOGRMYKMWITVTHSGELSNESFIFGRTPVPSVSHLGRSGNRNTTDSLWFNWSPASGD 1200  
Db 1141 FQNLLOGRMYKMWITVTHSGELSNESFIFGRTPVPSVSHLGRSGNRNTTDSLWFNWSPASGD 1200  
Qy 1201 FDFYELILYNPNGTKENWKDKOLTEWRFOQLVPGRYKYLWVTVTHSGDLNKNVTAESRTA 1260  
Db 1201 FDFYELILYNPNGTKENWKDKOLTEWRFOQLVPGRYKYLWVTVTHSGDLNKNVTAESRTA 1260  
Qy 1261 PSPSLMSFADIANTSIAITWKGPDWDYNDPELOWLPDALTVFNPYNNRSEGRIVY 1320  
Db 1261 PSPSLMSFADIANTSIAITWKGPDWDYNDPELOWLPDALTVFNPYNNRSEGRIVY 1320  
Qy 1321 GLRPGRSYQFNKTVSGDSWKTYSKPIFGSVRTKPKDIQNLHCRPQNSTAIACSWIPDPS 1380  
Db 1321 GLRPGRSYQFNKTVSGDSWKTYSKPIFGSVRTKPKDIQNLHCRPQNSTAIACSWIPDPS 1380  
Qy 1381 DFDGYSIECRMDTQVEVEFSRKEKESLLNIMMLVPHKRYLSIKVQSGAMTSEVVEDS 1440

Db 1381 DFDGYSIECRMDTQVEVEFSRKEKESLLNIMMLVPHKRYLSIKVQSGAMTSEVVEDS 1440  
Qy 1441 TITMIDRPPPPPHIRVNEKDVLSIKSSINFTVNCWFSFDTNGAVKYFTVVVRADSGDE 1500  
Db 1441 TITMIDRPPPPPHIRVNEKDVLSIKSSINFTVNCWFSFDTNGAVKYFTVVVRADSGDE 1500  
Qy 1501 LKPEQQHPLPSYLEYRHNASIRVYQNTYFASKCAENPNSNSKSFNIIKLGAEMLSGGKRD 1560  
Db 1501 LKPEQQHPLPSYLEYRHNASIRVYQNTYFASKCAENPNSNSKSFNIIKLGAEMLSGGKRD 1560  
Qy 1561 PTQOKFCODGPKPHTAYRISIRAFDTQDLDEKFKYPLYSDFPSLPITTESBPLFGAI 1620  
Db 1561 PTQOKFCODGPKPHTAYRISIRAFDTQDLDEKFKYPLYSDFPSLPITTESBPLFGAI 1620  
Qy 1621 EGVSAAGFLGLMVLAVVALLICROKVSNGRERSARLSIRDRPLSVHLNLGQKGRKTS 1680  
Db 1621 EGVSAAGFLGLMVLAVVALLICROKVSNGRERSARLSIRDRPLSVHLNLGQKGRKTS 1680  
Qy 1681 CPIKINQFEGHFMKLQADSNNYLLSKEYEELKDVGRNQSCDIALLPENRGNRYNNILPYD 1740  
Db 1681 CPIKINQFEGHFMKLQADSNNYLLSKEYEELKDVGRNQSCDIALLPENRGNRYNNILPYD 1740  
Qy 1741 ATRVKLSNVDDPCSDYINASYIPGNFRREYIVTQGPLPGTKDDFWKMWQNVHNVIM 1800  
Db 1741 ATRVKLSNVDDPCSDYINASYIPGNFRREYIVTQGPLPGTKDDFWKMWQNVHNVIM 1800  
Qy 1801 VTQCVKGRVKDHYMPADODSLYVYGDLLIQLMSESVLPWETIREFKICGEEQLDAHRLI 1860  
Db 1801 VTQCVKGRVKDHYMPADODSLYVYGDLLIQLMSESVLPWETIREFKICGEEQLDAHRLI 1860  
Qy 1861 RHFHYTVWPDHGVPEITQSLIQFVTRVDRYINRSPGAGPTVVHCSAGVGRGTGTFIALDRI 1920  
Db 1861 RHFHYTVWPDHGVPEITQSLIQFVTRVDRYINRSPGAGPTVVHCSAGVGRGTGTFIALDRI 1920  
Qy 1921 LQOLDSKSDSDIYGAVHDLRLHVRHVMVQTECOVYVYLHQCVRDVLARKLSEQENPLFPI 1980  
Db 1921 LQOLDSKSDSDIYGAVHDLRLHVRHVMVQTECOVYVYLHQCVRDVLARKLSEQENPLFPI 1980  
Qy 1981 YENVNPEYHREDPVYSRH 1997  
Db 1981 YENVNPEYHREDPVYSRH 1997

## RESULT 2

US-09-949-016-8049  
; Sequence 8049, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8049  
; LENGTH: 1903  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-8049

Query Match 94.4%; Score 9898; DB 2; Length 1903;  
Best Local Similarity 95.0%; Pred.No. 0;  
Matches 1897; Conservative 3; Mismatches 3; Indels 94; Gaps 2;  
Qy 1 MLSHGAGLALWITLSLQTLGABPERCNFTLAESKASSHSVSTQWIRILGSPCNFSLIYSS 60

Db 1 MLSHGAGLALMITLSLQGLAEPERCNFTLAESKASSHSVSIQWRILGSCNFSLIYSS 60  
Qy 61 DTLGAALCPTFRIDNTTYGCMQLDQAGTIYFNKIIISLDEBRTVVLQDTPPPARFGVSK 120  
Db 61 DTLGAALCPTFRIDNTTYGCMQLDQAGTIYFNKIIISLDEBRTVVLQDTPPPARFGVSK 120  
Qy 121 EKTSTSTGLHVMWTPSSGKVTSEYVQLFDENKQKOGVOIQBSTSNWETFFENLTAGSKYN 180  
Db 121 EKTSTSTGLHVMWTPSSGKVTSEYVQLFDENKQKOGVOIQBSTSNWETFFENLTAGSKYN 180  
Qy 181 IAITAVSGGKKSFSYTTNGSTVPSPVKDIGISTKANSLLISWSHSGGNVERYRLMLMDKG 240  
Db 181 IAITAVSGGKKSFSYTTNGSTVPSPVKDIGISTKANSLLISWSHSGGNVERYRLMLMDKG 240  
Qy 241 ILVHGGVVDKATSVAFHGLSPGYLYNLTVMTEAAGLQNYRWKLVRTAPMEVSNLKVTDN 300  
Db 241 ILVHGGVVDKATSVAFHGLTPGYLYNLTVMTEAAGLQNYRWKLVRTV- 288  
Qy 301 GSLTSLKVKQRPNGVDSYNIITLSHKGTIKESRVLAPWITETHFKELVPGRLYQVTVSC 360  
Db 289 ----- 288  
Qy 361 VSGELSAQMAVGRTPDKVANLEANNGRMRSLVWSMPPAGDWEQYRILLFNDVSVLL 420  
Db 289 -----PKVANLEANNGRMRSLVWSMPPAGDWEQYRILLFNDVSVLL 332  
Qy 421 NITVKEFTQYVMDTGLVPGQYVEVEVIVBSGNLKNSECCQGRTPVLAFLQLRVKHANE 480  
Db 333 NITVKEFTQYVMDTGLVPGQYVEVEVIVBSGNLKNSECCQGRTPVLAFLQLRVKHANE 392  
Qy 481 TSLIMWQTPVAWEKYYIISLADRDLLIHKSLSKDAKEFTFDLVPGRKYMATVTSISG 540  
Db 393 TSLIMWQTPVAWEKYYIISLADRDLLIHKSLSKDAKEFTFDLVPGRKYMATVTSISG 452  
Qy 541 DLKSSSVKGRTPVPAQVTDLHVANQGMTSLFTNWTQAGDVEFQVLLIHENVVIKNE 600  
Db 453 DLKSSSVKGRTPVPAQVTDLHVANQGMTSLFTNWTQAGDVEFQVLLIHENVVIKNE 512  
Qy 601 ISSETSRYSFHLKSGSLYSVVVTVTSGGISRQVVEGRTVPSSVSGVTNNSGRNDYL 660  
Db 513 ISSETSRYSFHLKSGSLYSVVVTVTSGGISRQVVEGRTVPSSVSGVTNNSGRNDYL 572  
Qy 661 SVSMLVAPGDVNDTEVTLSDHCKVQVSLVIAKSVRECSFSLTPGRLYTVTITRSGKYE 720  
Db 573 SVSMLVAPGDVNDTEVTLSDHCKVQVSLVIAKSVRECSFSLTPGRLYTVTITRSGKYE 632  
Qy 721 NHSFSQERTVPDKQGVSVNSARSXYLRVSWVHATGDFDHYEVIKNNFIOTKSI PK 780  
Db 633 NHSFSQERTVPDKQGVSVNSARSXYLRVSWVHATGDFDHYEVIKNNFIOTKSI PK 692  
Qy 781 SENECEVFLQVPGRLYSYVTVTKSQYBANQNGRTTPEPVKDLTLNRSTEDLHVTS 840  
Db 693 SENECEVFLQVPGRLYSYVTVTKSQYBANQNGRTTPEPVKDLTLNRSTEDLHVTS 752  
Qy 841 GANGDVOYETQLLFNDMKVPPPHLVNTATEYFTSLTPGRQYKILVLTISGDVQOSAF 900  
Db 753 GANGDVOYETQLLFNDMKVPPPHLVNTATEYFTSLTPGRQYKILVLTISGDVQOSAF 812  
Qy 901 LEGFTVPSAVKNIHISPNGATDSLTVNWTGCGDVDSYTVSAFHSQKVDSTQTPKHVFE 960  
Db 813 LEGFTVPSAVKNIHISPNGATDSLTVNWTGCGDVDSYTVSAFHSQKVDSTQTPKHVFE 872  
Qy 961 HTFHRLAGEQYQIMIASVSGSLKNQINNVGRTVPASVQGVIAADNAVSSYSLIVSWOKAA 1020  
Db 873 HTFHRLAGEQYQIMIASVSGSLKNQINNVGRTVPASVQGVIAADNAVSSYSLIVSWOKAA 932  
Qy 1021 GVAERYDILLTENGILLRNTSEPAATTKQHPEDLTPGKYKIQLITVSGGLFSKEAQTE 1080  
Db 933 GVAERYDILLTENGILLRNTSEPAATTKQHPEDLTPGKYKIQLITVSGGLFSKEAQTE 992  
Qy 1081 GRTVPAAVTDLRITENSTRHLSFRWTASEGELSWYNIFLYNDGCLQERAQVDPLVQSFS 1140

Db 993 GRTVPAAVTDLRITENSTRHLSFRWTASEGELSWYNIFLYNDGCLQERAQVDPLVQSFS 1052  
Qy 1141 FQNLQGRMYKMWIVTHSGELSNESFIFGRTPVASVHLRGSNRNTTDSLWFWNSPASGD 1200  
Db 1053 FQNLQGRMYKMWIVTHSGELSNESFIFGRTPVASVHLRGSNRNTTDSLWFWNSPASGD 1112  
Qy 1201 FDFVELLILYNPNGTKKENWKDKLTWRFOGLVPGRYKYLWVVTTHSGDLSNKTVAESRTA 1260  
Db 1113 FDFVELLILYNPNGTKKENWKDKLTWRFOGLVPGRYKYLWVVTTHSGDLSNKTVAESRTA 1172  
Qy 1261 PSPSLMSFADIANSTSLAITWKGPPDWTYNDPELQMLPRDALTVFNPYNNRKSSEGRIVY 1320  
Db 1173 PSPSLMSFADIANSTSLAITWKGPPDWTYNDPELQMLPRDALTVFNPYNNRKSSEGRIVY 1232  
Qy 1321 GLRGRSYQFNVKTVSGDSWKTYSKPIFGSVRTTPDKIONLHCRPQNSTAICSWIPDPS 1380  
Db 1233 GLRGRSYQFNVKTVSGDSWKTYSKPIFGSVRTTPDKIONLHCRPQNSTAICSWIPDPS 1292  
Qy 1381 DFDGYSIECRMDTOEVEFSRKLKLEKSLNIMMLVPHKRYLVSIVKQSGAGTSEVVEDS 1440  
Db 1293 DFDGYSIECRMDTOEVEFSRKLKLEKSLNIMMLVPHKRYLVSIVKQSGAGTSEVVEDS 1352  
Qy 1441 TITMIDRPPPPPHIRVNEKDVLSKSSINFTVNCNFSFSDTNGAVKYFTVVVREADGSDE 1500  
Db 1353 TITMIDRPPPPPHIRVNEKDVLSKSSINFTVNCNFSFSDTNGAVKYFTVVVREADGSDE 1412  
Qy 1501 LKPEQQHPLSYLEBYRNASIRVYQTYNYPASKCAENPNNSKSNFKILGAMESLGGKRD 1560  
Db 1413 LKPEQQHPLSYLEBYRNASIRVYQTYNYPASKCAENPNNSKSNFKILGAMESLGGKRD 1472  
Qy 1561 PTQOKFCGGLPKHTAYRISRAFTQFDEDLKFTKPLYSDFPSPITTESPLFGAI 1620  
Db 1473 PTQOKFCGGLPKHTAYRISRAFTQFDEDLKFTKPLYSDFPSPITTESPLFGAI 1532  
Qy 1621 EGVSAGLFLIGMLVAVVALLICROKVSCHGRERPSARLSIRDRPLSVHLNLGQK- 1587  
Db 1533 EGVSAGLFLIGMLVAVVALLICROKVSCHGRERPSARLSIRDRPLSVHLNLGQK- 1587  
Qy 1681 CPIKINGFEHFMKLQADSNYLLSKYEELKDVGRNOSCDIALLPENRGKRYNNILPYD 1740  
Db 1588 -PIKINGFEHFMKLQADSNYLLSKYEELKDVGRNOSCDIALLPENRGKRYNNILPYD 1646  
Qy 1741 ATRVKSNDVDDDCSDYINASYIPGNFRREYITVQGLPGTKDDFKWVWQVNHVLM 1800  
Db 1647 ATRVKSNDVDDDCSDYINASYIPGNFRREYITVQGLPGTKDDFKWVWQVNHVLM 1706  
Qy 1801 VTQCEVGRVKCDHYWADQDLSLYGDLILQMLSESVLPEWTIREFKICGSEQLDAHRLI 1860  
Db 1707 VTQCEVGRVKCDHYWADQDLSLYGDLILQMLSESVLPEWTIREFKICGSEQLDAHRLI 1766  
Qy 1861 RHFHYTVPDHGVPETTSQSLIQFVRTVRYNINRSPGAGTVPVHCSAGVGTGTFTIALDRI 1920  
Db 1767 RHFHYTVPDHGVPETTSQSLIQFVRTVRYNINRSPGAGTVPVHCSAGVGTGTFTIALDRI 1826  
Qy 1921 LQQLDSKSDVDIYCAVHDLRLHRVHMVQTECOYVYLHQCVRDVLRAKLRSEQNPLPPI 1980  
Db 1827 LQQLDSKSDVDIYCAVHDLRLHRVHMVQTECOYVYLHQCVRDVLRAKLRSEQNPLPPI 1886  
Qy 1981 YENVNPEYHRDVPYSRH 1997  
Db 1887 YENVNPEYHRDVPYSRH 1903

## RESULT 3

US-08-201-697-16  
; Sequence 16, Application US/08201697  
; Patent No. 5705623  
; GENERAL INFORMATION:  
; APPLICANT: Wiggins, Roger C.  
; APPLICANT: Thomas, Peedikayil E.  
; TITLE OF INVENTION: Mammalian Glomerular Epithelial Protein  
; TITLE OF INVENTION: 1  
; NUMBER OF SEQUENCES: 17

```

CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/201.697
; FILING DATE: 25-FEB-1994
; CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UM 9783
TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 278 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-201-697-16

Query Match 13.8%; Score 1444; DB 1; Length 278;
Best Local Similarity 95.7%; Pred. No. 3.5e-84;
Matches 266; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1682 PIKIQEGHPMKLQADSNLLSYKEYBELKDVGRNQSCDIALLPENRGKRYNNILPYDA 1741
Db 1 PIKIQEGHPMKLQADSNLLSYKEYBELKDVGRNQSCDIALLPENRGKRYNNILPYDA 60

Qy 1742 TRVLSNVDDPCSDYINASVPGNFRREYIVTQGPLPGTKDDFWKMWQNVNHNLMV 1801
Db 1 TRVLSNVDDPCSDYINASVPGNFRREYIVTQGPLPGTKDDFWKMWQNVNHNLMV 120

Qy 1802 TCQVEKGRVKCDHYWPAQDQSLYYGDLILQMLSESVLPWTIRFVKICGEEQDAHLIR 1861
Db 121 TCQVEKGRVKCDHYWPAQDQSLYYGDLILQMLSESVLPWTIRFVKICGEEQDAHLIR 180

Qy 1862 HFHYTVPDGHGVPETTQSLIQFVTRVDYINRSPGAGTVPVHCSAGVGRGTFTIALDRIL 1921
Db 181 HFHYTVPDGHGVPETTQSLIQFVTRVDYINRSPGAGTVPVHCSAGVGRGTFTIALDRIL 240

Qy 1922 QQLDSKDSVDIYGAVHDLRLHRVHMVQTECOYVYLHOC 1959
Db 241 QQLDSKDSVDIYGAVHDLRLHRVHMVQTECOYVICIHOC 278

RESULT 4
US-08-685-992-14
; Sequence 14, Application US/08685992
; Patent No. 5912138
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas
; APPLICANT: Flint, Andrew J.
; TITLE OF INVENTION: SUBSTRATE TRAPPING PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/201.697
; FILING DATE: 25-FEB-1994
; CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UM 9783
TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 278 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-201-697-16

Query Match 13.8%; Score 1444; DB 1; Length 278;
Best Local Similarity 95.7%; Pred. No. 3.5e-84;
Matches 266; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1682 PIKIQEGHPMKLQADSNLLSYKEYBELKDVGRNQSCDIALLPENRGKRYNNILPYDA 1741
Db 1 PIKIQEGHPMKLQADSNLLSYKEYBELKDVGRNQSCDIALLPENRGKRYNNILPYDA 60

Qy 1742 TRVLSNVDDPCSDYINASVPGNFRREYIVTQGPLPGTKDDFWKMWQNVNHNLMV 1801
Db 1 TRVLSNVDDPCSDYINASVPGNFRREYIVTQGPLPGTKDDFWKMWQNVNHNLMV 120

Qy 1802 TCQVEKGRVKCDHYWPAQDQSLYYGDLILQMLSESVLPWTIRFVKICGEEQDAHLIR 1861
Db 121 TCQVEKGRVKCDHYWPAQDQSLYYGDLILQMLSESVLPWTIRFVKICGEEQDAHLIR 180

Qy 1862 HFHYTVPDGHGVPETTQSLIQFVTRVDYINRSPGAGTVPVHCSAGVGRGTFTIALDRIL 1921
Db 181 HFHYTVPDGHGVPETTQSLIQFVTRVDYINRSPGAGTVPVHCSAGVGRGTFTIALDRIL 240

Qy 1922 QQLDSKDSVDIYGAVHDLRLHRVHMVQTECOYVYLHOC 1959
Db 241 QQLDSKDSVDIYGAVHDLRLHRVHMVQTECOYVICIHOC 278

RESULT 5
US-09-144-925-14
; Sequence 14, Application US/09144925
; Patent No. 5951979
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas
; APPLICANT: Flint, Andrew J.
; TITLE OF INVENTION: SUBSTRATE TRAPPING PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02421-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
```

```

; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/237,940
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael P.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27866/31954
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1337 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-854-585-2

Query Match      11.9%; Score 1247; DB 2; Length 1337;
Best Local Similarity 28.9%; Pred. No. 1.3e-70;
Matches 417; Conservative 215; Mismatches 547; Indels 264; Gaps 61;

QY    641 TYPSSVSGVTNNNGRNDYLVSWLVA PGDVVDNYEVTLSHDKVQVSLVIKSVRECSFS 700
      || :|| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db    50 TVATGENGITQISSTAESFHKKQTGP---QVETNTSEDG---ESSGANDSLR----- 97

QY    701 SLTPGLRYLTVITTRSGKYENHSFSOE---RTVPDKV---QGVSYSNSARSDDLRLRVSWVH 754
      || :|| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db    98 -----TPEGNGTDGASOKTSSPGSPFDIKAVSIPT-----NVILTWS 141

QY    755 ATGDFDHYEVTIKNKNFNFIQTAKSIKSENE-----CVFQLVPGRLYSVTVTTKS 804
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db    142 NDTASEKYVVKH-----KMENEKITVVHQPCWNI TGLRPATSVFSITPCI 190

QY    805 GQYEANEQ--GNGRTI-----PEPVKDLTLRNKRSTEDLHVTVWGANG-----DVDQY 849
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db    191 G---NETWGDPRIKVITEPIPVSDLRVAHGCEGCSLSWNGNGTASCRCVLLESIGSH 246

QY    850 ETQLLFNDMKVPPPHLVNATEYEFTSLTPGRQYKILVITI-----SGD--VQSAPFIE 902
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db    247 E--ELTQDSRL-----QVNISDLKPGVQVINPYLQSNKTGKDPLAQKVAW-- 291

QY    903 GFTVPSAVKNIHISPNGADSLTNVMTPGGGDVDSYTVSAFRHSQKVDQSOTIPKHFVEHT 962
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db    292 ---MFAIQBE---ACGGAPPPLCM-MSPFVGVPDPSS-----GQQRUTEVL----- 331

QY    963 FHRLBAGEQQYQMIAAS--VSGSLKQNINNVGRTVPASPVGQIVTADNAYSYSIIVSQKAA 1020
      || :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

```

QY	1021	6VAE---	RYDILL	LTENG	ILLANT	SE	PATT	KQHK	PED	TPGK	----	YKIQ	ILTV	SGGLF	1073
Db	391	NSSSN	YTKI	HVAG	ETDS	SNLN	SEPA	-----	VIP	GR	SR	STF	YNIT	VC	PLGDIE 442
QY	1074	SKEA	TEGR	TPAA	VTDL	RI	TENS	TRHL	SFRW	TASE	GL	SWNI	FLYN	PDGN	LQRAQVD 1133
Db	443	GT	PG	FLQ	HTP	VP	VSDF	RVTV	VS	TE	IG	LAW	SSHDA	ES	FOWHI - - - - - TOEGAGN
QY	1134	PLVQ	SF	SFQ	NLLQ	GR	MYK	WVI	VTH	--	S	GELS	NES	I	FOR
Db	499	TTNQ	II	IG	LP	PG	KY	CFE	I	VP	K	PG	NG	T	EGAS
QY	1192	FNWS	PAS	G	D	F	D	F	-	YEL	LY	N	P	NG	T
Db	558	L	D	W	K	S	P	D	G	A	S	E	V	Y	H
QY	1248	DLS	NK	VT	A	E	S	R	T	A	P	S	P	P	S
Db	615	G	D	P	N	S	T	A	Q	Y	T	R	-----	P	N
QY	1308	PYNN	RK	S	E	G	R	I	Y	V	G	-----	L	R	P

Db 663 KAGSSNATQVVTIDGITATVTELPGSSYTVLFAQVGDGKSL-EPGRKSFCTDPAS 721  
Qy 1358 IONLHCR-PQNSTALACSWIPDDDFDCYSIE-----CRKMDTQVEFSRLEKEKS 1408  
Db 722 MASDFCEVVPKPAVLVLTCPGAGNAGFELEVSSGAWNATHLESSENGTETRTVT 781  
Qy 1409 LLNIMMLVPHKRYLVLSIKVQAGMTSEYVEDSTITMIDRPPPP--PPHVRNKRDLVLISK 1466  
Db 782 YLNF-----STSYNISITTVSCGKMAAPTRNTCTTGIDPPPPDGSPI-----TSVSH 830  
Qy 1467 SSINFVNCWSPDNGAVKFTVVVREADSGDELKPEQQHPLPSYLEYRHNAIRVYQT 1526  
Db 831 NSVK--VKFSFEASHGPIKAYAVITTG-----EAGHPSADVLKTYTDFKKGASD 880  
Qy 1527 NY--FASKCAENPNSNKS-----FNILGAMESISGGKRDPTQKFCGDKPLKPTAVRIS 1580  
Db 881 TYVILIRTEKGRSQSLEVLKYEIDVGNESSTLG-----YNGKLEPLGYSRAC 931  
Qy 1581 IRAFTQ-----LFD--BDLKEFTKPLYSDFSLPITTESEP--LFGAIEGVSAGL 1627  
Db 932 VAGFTNITFHPQKGLIDGAEVYSFSR--YSDA-VSLP---QDPGVICGAVFGCIFA 984  
Qy 1628 FLIGMLVAVALLICROKVSGRHRPBARLSIRDRPLSVHLNLGQGNKRTSPKIKNQ 1687  
Db 985 LVI---VTVGFI FWRKK--RKDAKNEVFSQIKP-----KSKLIRVEN 1025  
Qy 1688 FEGHFMKLOADSNYLLSKYEELKDVGRNOSCDIALLPENRGNKRYNNILPYDATRVKLS 1747  
Db 1026 FEATFKQQADNSGCFABEYEDLKLVLGSLQPKYAEALAEKNRGNRYNNVLPYDLSRVKLS 1085  
Qy 1748 NVDDPCSDYINASYIPGNFRREYIVTQGPLTKDDFMQVWQONVHNITVMVTOCVEK 1807  
Db 1086 -VQTHSTDDYINANYPGYSKDFIATQGPLNTLXDFWVWKEKNVYALIMLTCKVEQ 1144  
Qy 1808 GRVCDHYWADQSLYGLIOLMSVLPENTIREFKICGEEQIDAHLRIKHFTYV 1867  
Db 1145 GRKCEBYWPSKQ-AQYGDITVAMTSEIVLPENTIRDFVKNITQTSSEHPL-RQFHFTS 1202  
Qy 1868 WPDHGVPTTOSLQFVRTVDYINRSPGAPTVVHCSAGVGTGTFIADRLILQQLDSK 1927  
Db 1203 WPDHGVPTDOLLNFRYLVRYDMKQSPSPILVHCSAGVGTGTFIADRLILQIENE 1262  
Qy 1928 DSVDIYGAHDLRLHRYVHMVQTEQYVYLHQVCRDVLRLARKLS-----EQENPLPFIYEN 1983  
Db 1263 NTVDYGVIVDLMRHPMLVQTEQYVFLAQCVLDIVRSQKSDVLIYQNTAMTIYEN 1322  
Qy 1984 VNP 1986  
Db 1323 LAP 1325

## RESULT 7

US-09-447-533-2

; Sequence 2, Application US/09447533

; Patent No. 6552169

; GENERAL INFORMATION:

; APPLICANT: Tonks, Nicholas K.

; ; Oetman, Arne

; TITLE OF INVENTION: DENSITY ENHANCED PROTEIN TYROSINE

; PHOSPHATASES

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Seed IP Law Group PLLC

; STREET: Suite 6300, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/447.533  
; FILING DATE: 23-No. 6552169-1999  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rosenman Ph.D., Stephen J.  
; REGISTRATION NUMBER: 43,058  
; REFERENCE/DOCKET NUMBER: 200125.402C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1337 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-447-533-2

Query Match 11.98; Score 1247; DB 2; Length 1337;  
Best Local Similarity 28.94; Pred. No. 1.3e-70;  
Matches 417; Conservative 215; Mismatches 547; Indels 264; Gaps 61;  
Qy 641 TVPSSVSGVTWNNSGRNDYLSVWLVAPGDVNDYEVTLSDHGKVVQSLVIKAVSRECSFS 700  
Db 50 TVATGENGIITQISSTAESFKHQNGTGP-----QVETNWTSEDG-----ESSGANDSLR----- 97  
Qy 701 SLTPGRLYTVITTRSKYENHHSFSQE---RTVPDKV---CGVSVNSASRDYLRVSVWH 754  
Db 98 -----TPEQSGNCTDGAOSKTPSPSTGSPVFDIKAVSISPT---NVILTWS 141  
Qy 755 ATGDFDHEVETIKNNFIQTKSPIKSENE-----CVFVQLVGRLYSVTVITKS 804  
Db 142 NDTAAASEYKVVXK-----KXENETITVHQPCWNCITGLRPATSYVPSITPGI 190  
Qy 805 GOYEANEQ-GNGRTI-----PEPVKOLTLNRSTEDLHVTWVGANG-----DVEDQY 849  
Db 191 G-----NETWGDPRVIVTEPIPVSDLRVAHGCEGCSLSWNGNGTASCRLVLESIGSH 246  
Qy 850 EIQLLFNDKMFPHLVNTATYRFTSLPPGQYKILVLT-----SGD--VQOSAFIE 902  
Db 247 E--ELTQDSRL-----QVNISDLKPGVQYINPYLLOSNKTKGDPKQKVAW-- 291  
Qy 903 GFTVPASVAKNIHSPNGATDSLTVNTPGGGDVDSYTVSAFRHSQKVDSTIPKHVPEHT 962  
Db 292 ---MPAIQRE---AGQAGPPPLCM-MSPFVGPVDPSS-----GQQSRDTEVL----- 331  
Qy 963 FHRLEAGEQYQIMIAS--VSGSLKNQINVVGRTPASVQGVIAADNAYSSYSLIVSQKAA 1020  
Db 332 LVGLEPGRTRNATVYSOANGTEGQQAIEFRINAIOVFDVTAVN-ISAATLTLIKVSD 390  
Qy 1021 GVAE---RYDILLTENGILLRNTSEPAATKQHKFEDLTGPKK-----YKIOILTIVSGGLF 1073  
Db 391 NESSSNVYKIHVAGETDSSNLNVSEPR-----VIPGLRSSTFYNTITVCPVLGDIE 442  
Qy 1074 SKEAQTEGRTPAAVTDLRITENSTRHLSRFTWASSEGELSWYNIFLNPNQNLQERQVD 1133  
Db 443 GTPGFLQVHTPPVPVSDRFTVTVVSTTEIGLAWSHDABSQMH-----TQSGAGNSRVEI- 498  
Qy 1134 PLVQSFQNLQGRMYKMWIVTH--SGELSNESFIPGRTPVPSVSHLGRSNRNTTDSLW 1191  
Db 499 TTWQSIILIGLFPGTKYKCFEIVKPGNGTSGASRTVCNRTVPSAVFDHVVVYVTTT-MW 557  
Qy 1192 FNVSPASGDPDF-YELILYNPNTKKENWKDKLTWRFOGLVPGRYKVLWV---VTHSG 1247  
Db 558 LDWKSPPDGAEEYVHLVIESKHGNSHTSTYDKAIT---LQGLIPGLTYNITISPEVDHVW 614  
Qy 1248 DLSNKVTAESRTAPSPSLMSFADIANSLAITWKGPDPDWDYDNFELQWLPRDALTVFN 1307  
Db 615 GDPNSTAQYTR-----PSNVNIDVSTNTTAATL-----SWQNFDDASPTV---SYCLLIE 662  
Qy 1308 PYNRRKSEGRIVYG-----LAPGRSYQFNVKTVSGDSWKTYSKPIFGSVRTKPKD 1357



Db 663 KAGNSSNATQVVDIGITDAITVTELPFGSSVTVLFAQVGGIGKSL-EPGRKSCFDPAS 721  
Qy 1358 IONLHCR-PONSTAIACSWIPDSDFDGYSIE-----CRKMDTOVEPSRKLKKEKS 1408  
Db 722 MASPDCEVVPKPALVLTCPGANAGFELEVSSGAWNATHLESCSSENGTEYRTEVT 781  
Qy 1409 LLNIMMLVPHKRYLVSVKQAGMTSEVVEDSTLTMDRPPPP--PHIRVNEKDVLSK 1466  
Db 782 YLNF-----STYSNISITTVSCGMAAPRNTCTTGITDPPPPGSGNI-----TSVSH 830  
Qy 1467 SSINFTVNCWSFSDTNGAVKFTVVVREADSGDELKPEQHPSPSYLEYRHNASIRVVQT 1526  
Db 831 NSVK--VKFSGEASHGPIKAYAVILTTG-----EAGHPSADVLKYTYDDPKGASD 880  
Qy 1527 NY--PASKCANPNNSKS-----FNKLGAMESLGGKRDPTQKFCGPKLPHAYRIS 1580  
Db 881 TYVTYLIRTEKGRSQSLSEVLKYEIDVGNSTTLG-----YNGKLEPLGSYRAC 931  
Qy 1581 IRAETQ-----LFD--EDLKEFTKPLYSDFPSLPITTESEP--LFGAIEGVAGL 1627  
Db 932 VAGFTNITFHQNKGLIDGABSYVSFSR--YSDA-VSLP-----QDPGVICGAVFCIFGA 984  
Qy 1628 FLIGMLVAVALLICROKVSHERPERSARLIRDRPLSVHLNLGQGNKRTSCPIKINQ 1687  
Db 985 LVI---VTVGCFIWRKK--RKDAKNEVSFSQIKP-----KSKLIIRVEN 1025  
Qy 1688 FEGHMKLOADNSYLLSKYBELKDVGRNQSCDIALLPENRGNKRYNNILPYDATTRVLS 1747  
Db 1026 FEAVFKQADNSCNGFAEYEDLKLIGSQPKYAAELAEANRGNKRYNNVLPYDISRVKLS 1085  
Qy 1748 NVDDPCSDYINASYIPGNFRERYIVTQPLGPKTODPFKMWQWQVNHVIMVTCQVEK 1807  
Db 1086 -VOTHSTDDYINANYMPGYSKSDFIATQGPLNLTDFWNRWKNVYAILMLTKCEVQ 1144  
Qy 1808 GRVKCDHWPADQSLYVGLILOMLGESVPEWTIREFKICGEOQLDAHRLIRHFHTV 1867  
Db 1145 GRTICEYWPBQK-AQYGDITVAMTSEIVLPEWTIIRDFTVKNITQTSSEHPL-RQHFST 1202  
Qy 1868 WPDHGVPTTQSLQFVTRVDRYINRSPGAGPTVHCSAGVGRGTFTIALDRILQQLDSK 1927  
Db 1203 WPDHGVPTTDLINFRVLYRDMYKQSPBSPILVHCSAGVGRGTFTIALDRILYQIENE 1262  
Qy 1928 DSVDIYGAHDLRLHRVHMVQTEQYVYLHOCVDVLRARKLS-----EQBNPLFIYEN 1983  
Db 1263 NTVDVYGVYDLRMRPLMVQTEQYVFLNQVLDIVRSQKSDVLDIYQNTTAMTIYEN 1322  
Qy 1984 VNP 1986  
Db 1323 LAP 1325

## RESULT 8

PCT-US95-05512-2  
; Sequence 2, Application PC/TUS9505512  
; GENERAL INFORMATION:  
; APPLICANT: Tonks, Nicholas K. and stman, Arne  
; TITLE OF INVENTION: Density Enhanced Protein Tyrosine  
; TITLE OF INVENTION: Phosphatase  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Borun  
; STREET: 233 South Wacker Drive, Suite 6300  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/05512  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Borun, Michael F.  
; REGISTRATION NUMBER: 25,447  
; REFERENCE/DOCKET NUMBER: 27866/31954  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1337 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US95-05512-2  
  
Query Match 11.9% Score 1247; DB 4; Length 1337;  
Best Local Similarity 28.9%; Pred. No. 1.3e-70;  
Matches 417; Conservative 215; Mismatches 547; Indels 264; Gaps 61;  
  
Qy 641 TVPSSGVTVNNSGRNDYLSVWLVAPGDVDNYEVLTHDGVVQSLVIKSVRECSFS 700  
Db 50 TVATGENGIIQISSTAESFHKQGTGP---QVETWTSDEG---ESSGANDSLR----- 97  
Qy 701 SLTPGRLYTVTITRISKYENHFSQEB---RTVPDKV---QGVSVNSARSDDYLRVSWH 754  
Db 98 -----TPEQSGNGTGDGASQKSTGSPVFDIKAVSISPT---NVILTWS 141  
Qy 755 ATGDPFDHYEVTIKNNFIQTKSPIKSENE-----CVFQVLVPGRLYSVTVTTKS 804  
Db 142 NDTAAASEYKVVXK-----KXENKXITTVVHPWCNITGLRPAISYVSIPTGI 190  
Qy 805 GQYEANEQ-GNGRTI-----PEPVKDLTLNRSTEDLHVTWSGANG-----DVDQY 849  
Db 191 G-----NETWGDPRVIVTEPIPVSDLRVAHGCEGCSLSWSNGNGTASCRLVLESIGSH 246  
Qy 850 EIQLLFNDKMFPPFHLVNTATEYRFTSLTPGRQYKILVLT-----SGD--VQOSAFIE 902  
Db 247 E--ELTQDSRL-----QVNIIDLKPGVQYNNINPYLLQSNKTKGDPPLAQKAV-- 291  
Qy 903 GFTVPSAVKNIHISPGATDSLTVNWTGPGGDVDSYTSAPRHSQKSDVSTQIPKHVEHT 962  
Db 292 ---MPAIORE---AGQAGAPPLCM--MSFPVGPVDPSS-----GQOSRDETVL----- 331  
Qy 963 FHRLEAGEQYQIMIAS--VSGSLKNQINNVGRTVPASVQGVADNAYSSYSLIVSWOKAA 1020  
Db 332 LVGLEPGRTRNATYVSOAANGTEGQPOAIEFRNATVAVN-ISAATSLTLIKVSD 390  
Qy 1021 GVAB---RYDILLTENGILLRNTSEPAITKQHKFEDLTGKK-----YKIQILTVSGGLF 1073  
Db 391 NESSSNYTKIHVAGETDSSNLNVSEPA-----VIFGLRSSTFYNTVCPVLGDIE 442  
Qy 1074 SKAQTEGRTVPAVTDLRLITENSTHLSFRWTASEGELSWYNIFLYNDGNLQRAQVD 1133  
Db 443 GTPGFLQVHTPPVPSVDFRVTVSTTEIGLAWSHDAESFQMH-----TOEGAGNSRVEI- 498  
Qy 1134 PLVQSPSFQNLQGRMYKWIIVTH--SGELSNESFIFGRTPVPSVSHLGRNRTTDSLW 1191  
Db 499 TTNQSIILIGLFPCTKYCFEIVPKGNGTEGARTVCNRTVPASVDFIHHVYVVTTE-MW 557  
Qy 1192 FNMSPASGDFDP-YELILYNPNGTKENWKDOLTEWRFOGLVPRKYVLWV-----VTHSG 1247  
Db 558 LDMKSPDGASEYVYVHLVIESKHGSHNTSTYDKAIT---LQGLIPGLTYNITISPEVDHW 614  
Qy 1248 DLSNKVTABSRTPSPPLMSFADIANSTLAITWKGPDDTDYNDYDFLOWLPRDLALTFN 1307  
Db 615 GDPNSTAQYTR-----PSNVSNIDVSTNTTAATL-----SWQNFDDASPTY---SYCLLIE 662  
Qy 1308 PYNRKSEGRIVYG-----LRPGRSVQFNKVTSGDSWKTSYKPIFGSVRTKPKD 1357  
Db 663 KAGNSSNATQVVDIGITDAITVTELPFGSSVTVLFAQVGGIGKSL-EPGRKSCFDPAS 721







Db 604 TIYATELTDNRAFQMTTVDNSFLITGLKKYTRYKMRVAASHTVGESSLSBENDIFVTRPE 663  
Qy 394 -----LVVSPPA---GDWEQYRILLPNDVSVLLNITVQKBEQY 431  
Db 664 DEPESSQDVQVGTGVPSELRLKWSPPKPKNGIIIAEYVQLQNDTLFVKNT---STTDI 720  
Qy 432 VMDDTGLVPGQYEVYEVESGNLKNSERCQGRTVPLAVQLQRLVKHANETS--LSIMWQT 489  
Db 721 IISD--LXPYLYNISI-----RS-----YTRLGHGNQSSLLSVRTSE 757  
Qy 490 PVAWEKYIISLADRDLLIHKSLSKDAKEFTFDLVPRGRYMATVTSISGDLKSSSVK 549  
Db 758 TVPD-----SAPEN---ITYKNISSGEIISF---LPPRSPNGIIQYTYILKRSNSHE 805  
Qy 550 GRTVPAQVTDLHVANQMTSSLFNTWTAQGDVEFYQVLLHENVVKNESISSETSRYS 609  
Db 806 ARTINT-----TSLTQITGLKKYHYVI----- 829  
Qy 610 FHSLSKSGSLYSVVVTVTSVG--GISSROV--VVBGRTPVPS-----VSGVTVANSGRND 658  
Db 830 -----EVSASTLKGEGIRRPISILTEEDAPDPPQNFVSKQLSGVT----- 872  
Qy 659 YLSVSWL--VAPGDVDNVEVTLSDHGKVVQSLVIKSVRECS--FSSILTPGRLY--TWTI 712  
Db 873 --MLSQPPEPENGILLYYTVVWD---KSLRALINATEASLVLSLDYNDVYGACVTA 926  
Qy 713 TTRSGKYENHGF-----SOERTVPDKVGVSVNSARSDYLRVSW---VHATGDFDH 761  
Db 927 STRFGDGNARSSIIINFRPEGEPSDPPNDVHVNLSSS-----IILFWTPVKPKNGIIQY 982  
Qy 762 YEVITKKN-----NFIOTKSIKPSSENECFVQVLPVGLYS-----VTVTKSQYEANE 811  
Db 983 YSVYQNTSGTFVQNFLLQVTKSSDNVTSVARIYLAIFSYTFWLTASTSVG--NGNK 1040  
Qy 812 QGN-----GRTIPE-PVKDLTLNRSTEDLHVW---SGANGVDVDOYEIQLLFNDMKVF 861  
Db 1041 SSDIHHVTDODIEGPGVGNLTFSISSTALHVSWEPPSQNGLVFY-----LSNLQOS 1096  
Qy 862 PPFLHVNATYR-----FTSLTPGRQYKILVLTISGDVQSAFIEGFT-----VP 907  
Db 1097 PPREHMPPLVTVENSIDPDDLEKYTDY-IPKITS---TEKGFSETVTLQHLKHTEDVP 1152  
Qy 908 SAVKNIHSPNGATDSLTVNWP---GGGDVDSVTSVAFHSQKVDQOTIPKHVFEHTFH 964  
Db 1153 DTPPIINTFNKLSSTILLSWDDPLKPKNGAILGYH-----TLQPHANHTF- 1199  
Qy 965 RLEAGEQYQIMIASVSGSLKQINVGRTV---PASVOGVADNA-----Y 1007  
Db 1200 -VTSGN--HIVLEBSPFLYSFPAARTMKGLGFSSILFFYTBESAPLAPPQNLINY 1256  
Qy 1008 SSVSLIVSWOKA---AGVAERYDILLTE--NGILLRNTSEPATYKQHPEDLTPGKKYK 1062  
Db 1257 TSDFWLWTSPLSPGGIVKVYSFKIHEHETDTVFYKXNISLQT--DAKLEGLEPVSYS 1314  
Qy 1063 IQILTV-----SGGLFSKEAQ--TEGRTVPAATVDLRIENSTRHLSFRW---TASEGELSW 1114  
Db 1315 VVSFAFTKVGNGNOYSNVVEFTQESPEAVRNEICVARDWQSVSVRWDPPRKTNGIIH 1374  
Qy 1115 YNIFLYPDGNLQERAQVPLVQSFQNLQGRMYKWI--VTHSGELSNES---FIF 1168  
Db 1375 YMITV---GGN---STKVSPPDPTFTYKLPNTSYVFEVRASTASAGE--GNESKCDISTL 1427  
Qy 1169 GRTVPASVSHLRGSRNRTTSLWFWNSPASGDFEYB--LILYNPNGTKKENW----- 1219  
Db 1428 PETVPSAPTNVAFSNVQST--SATLTWTKPDITFGYFQNYKITTLQRAQCKREWEPEBCIE 1486  
Qy 1220 --KOKDL-----TWRFGVLVPGKRYLWV--THSGDLSKNVTAERTAPSP--SLMSF 1269  
Db 1487 HOKQOYLYEANTQETVHGLKKFRWYRQVAASTNVGSNASEWISTQTLPGPDGPPEN 1546  
Qy 1270 ADIANTS---LAITWKGPPDWT-----DYNDFELOWLPRDALTVFNPNRKS 1314  
Db 1547 VHVATSPFGINISWSEPAVITGPTFYLLIDVKSVDODDFFNISFLK-----SNEEN 1596

Qy 1315 EGRIVYGLRPSRSQFNVKTVSGDSWKYTK-----PIFGSVTKPKDKIQNLHCRP 1365  
Db 1597 KTTEINNLEVTRYSVVITAFVGNVSRAYTDGKSSABEVIIITLESVPKDPNNMTFOKIP 1656  
Qy 1366 QNSTAIACSWIPDSDDFDQ-----YSIECKMDTOEVE---FSRKLKESKSLNIMM-LV 1416  
Db 1657 DEVTKFOLTFLPP--SQPNGNIRVYQALVYREDDTFAVQIHNFSLIQTDTYSIIAMLSGLK 1715  
Qy 1417 PHKRYLVI--KVOSAGMTSEVEDSTITM-----IDRPPPPPHIRVAKNOVLISKSSINF 1471  
Db 1716 GGHTYNIISVAINASAGAPKV--QMRITMDIKAPAPKSKPIPIRDATGKLLTSTTITI 1773  
Qy 1472 TVNCSWESDINGAVKYFTVVVREADSGDELKPEQOHPPLSVLEYVRHNASIRVQTN--YP 1529  
Db 1774 RMPICYNDHDPGRINNVQVLAETGA-----QODGNVTWKYDAYFNKA-RPYFTNREGFP 1826  
Qy 1530 ASKCAENPNNSKSFNI--KLGAEMESLGKGRDPTQOKFCQDGLPKPHPTAYRISIRAFATQLF 1598  
Db 1827 NPPIEGKTKFSGNEEIVVIGADNACM---IPGNEEKICNGPLKPKQYLFKFA----- 1878  
Qy 1589 DEDLEKFTKPLYSDTFPSLPITTESEPLF-CAIEGV-SAGFLIGMLV---AVVALLICR 1643  
Db 1879 TNVWGQFTDSEYSD-----PIKTLGEGLSERTVEIILSVTLICILSIILLGTAFAFVRIR 1933  
Qy 1644 QKVSHG-----RERPSARLSIRDRPLSV-HLNLGOK-----GNRTSCPIKINOPEGH 1691  
Db 1934 QOKEGGTYSPRDAIITDTFKLDQTLTADLEKOLBELTLLSYRKSIRKPSKKSFLQH 1993  
Qy 1692 FMKLOADSNVLLSKEYBELXDVGRNOSCDIALPENGRKRYNNILPVDATRVKLSNVDD 1751  
Db 1994 VBEICTNSNLKQGEFSELFPKFLQDLSSTDADLPWNAKARFPNIKPYNNRVKLIADVS 2053  
Qy 1752 DPCSDYINASYIIPGNFRREYIVTQGLPGTKODFWQVWQVONVHNIVMTVCYKGRVK 1811  
Db 2054 LPGSDYINASVSVGLPCNFEIATQGLPGTVGDFWQVWQVONVHNIVMTVCYKGRVK 2113  
Qy 1812 CDHYWADQDSL--YGDLLIQLMSSESVLPWTWTFREFKICGEEQLDAHLRIHFHYTWPD 1870  
Db 2114 CHQWPDNDKPVTVFGDIVITKLMEDIQIDWTIRDLKI---ERHGDQWTVRQCNTGWE 2170  
Qy 1871 HGVPETQSLIQVTVTVRVDYINBSPAGPTVHCSAGVGRGTPIADRLILQQLDSKDSV 1930  
Db 2171 HGVPENTTPIHFVKLVKVR--TSRAHDTTPMVHCSAGVGRGTGVFIADHLHTQHINNHDV 2228  
Qy 1931 DIYGAVHDLRLHRVHMVQTECQVYVYLHCQVRDVL 1964  
Db 2229 DIYGLVAELSERMCMVQNLQAQIFLHCQILDLL 2262

## RESULT 12

US-08-201-697-4  
; Sequence 4, Application US/08201697  
; Patent No. 5705623  
; GENERAL INFORMATION:  
; APPLICANT: Wiggins, Roger C.  
; APPLICANT: Thomas, Peedikayil E.  
; TITLE OF INVENTION: Mammalian Glomerular Epithelial Protein  
; TITLE OF INVENTION: 1  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

Qy	1251	NKVTASRTAPSPPSLMSFADIIETSLAIWKGPDPDWDYNDTE--LQWLPRDALTVENPY	1309
Db	622	DSSTISPIITAPVAPEITS--VEYENSLIYISWTYGGDDTTDLSHGRMLHWM-----	669
Qy	1310	NNRKSEGRIVYGLRPGRESYQFNKTVSGDSWKYTKSPKIPFOSVPT-----KPDKIQNLHCR	1364
Db	670	-----VVAEGKKIKKSVTRNVNVTAILSLPPGIYNL-----	701
Qy	1365	PONSTAIACSWIPPDSDFDGYSIECRKMDTQVEFEKSKLEKESKLLINMLVPHKRYLVS	1424
Db	702	-----SVTACT-----ERGSNTSMRLVKL-----	721
Qy	1425	IKVQSAGMTSEVVEDSTITMIDRPPPPPPHIRVNEKDVLSKSSINPTVNCWSFSDTNGA	1484
Db	722	-----EPAPKSLFVANKTQ-----SVTLWVEE--GV	748
Qy	1485	VKYFTVVVREADSGDELKPPQOHPHLPSELYEYRNINASIRVVQTNYPASKCAENPNSKSF	1544
Db	749	ADFFEVPCQVGSSQKTKLQEPVAVSS-----HWVTI-----	780
Qy	1545	NIKLGAEMESLGGKRPDTQOKFCGGLPKPHTATYRISIRATQOLFDEDLKEFTKPLYSDTF	1604
Db	781	-----SLLPATAVNCSTVSPSH-----DSFSV--PTF	806
Qy	1605	FSL--PITTESEPLFGAIEGVS--AGLFLIGMLVAVALLICROK--VSHGRERPSARL----	1657
Db	807	IAYSTWTEWNNPVNIVSLAILSTLLIGLL--LVTLILRKHLQWARECGAGTFWNFA	864
Qy	1658	SIRDRPLSVHMLGQGNKRTSCPIKIQFEGHFKMLQADSNLYLSKEYEELKDVGRNQ	1717
Db	865	SLERDGLKPLYNKSKGLKKRKLTPVQLDDFDAYIKDMAKSDYKFSLOPEELKLGLDI	924
Qy	1718	SCDIALLPENRGKRNYNILPYDATVKLSNVDDDDPCSDYINASYPGNFRFEYIVTQG	1777
Db	925	PHFAADPLNRCNRYNLPYDFSRVLVSMNEEBGADYINANYIPGYNPSOEYIATQG	984
Qy	1778	PLPGTKDDDFWKVWEQNVHNIVMVTOCVEKGRVKCDHYWPADQDSLYVYGDLIIQLMLESV	1837
Db	985	PLPETRNDFWKVLVQKSKQIIVMLTOCNEKGRVKCDHYWPFTEPTAYGDIITVEMISEE	1044
Qy	1838	LPETWIREFKICEBQDLAHLRHRPHYTVWPDHGV--ETTQSLIQFVRTVRDYINRSP	1895
Db	1045	QDDWACKHPRI---NYADEMODVMHENYTAWPDHGVPTANAABSIILQFVHMVRQQA	1100
Qy	1896	GAGPTVHVSAGVGRGTFTIALDRILQOLDSKDSVDIYGAHDLRLHRVHMVQTECQYVY	1955
Db	1101	-KGPMIHCBSAGVGRGTFTIALDRLLQHIRDHFVDILGLVSEMRSYRMSWQTEBQYIF	1159
Qy	1956	LHCVRDVLRLARKLSEQENPLPIYENVN	1985
Db	1160	IHCQVLMWMKKK---QQICISDVIYENVNS	1186

RESULT 13

US-09-949-016-8051

; Sequence 8051, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8051  
; LENGTH: 1246  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-8051

Query Match 8.7%; Score 913.5; DB 2; Length 1246;  
Best Local Similarity 24.0%; Pred. No. 2.1e-49;  
Matches 367; Conservative 199; Mismatches 483; Indels 481; Gaps 63;

Qy 571 LFTNWTQAGDVEFYQVLLIHENVIKNES-----ISSETSRY----- 608  
Db 81 LFKNAT-----AFHVTVDNNIVVLEASDVSPASVYVVKITGESKNYFFEEFENS 134  
Qy 609 -----SFHLSKSGSLSVVTVVSGGI-----SSRQVVVEGRTPVSSVSGVTNNNG 655  
Db 135 TLPPVPVIFKASYH-----GLXYIITLVVNGVNVTKPSRSITV-----LTKPLPVTSVSDYK 188  
Qy 656 RNDYLSVSWLAPGDVDNY--EVTLSH-DGKVQSLVI-----AKSVRECSFSSILTCRL 707  
Db 189 PSPETGVLFIEHYPEKYNVFRVNIISYWEKDFRTMLYKDFKGTV-----FNHWLPGMC 244  
Qy 708 YT-----VITTRSGKYENHSFSQERTVPDKVQGVSVNSARSVDYLR--- 750  
Db 245 YSNITPQLVSEATFNKSLVEYSG--VSHEPKQHRTAPYPQNISV-----RIVNL 293  
Qy 751 ---SWHATGDFDHYEVTIKNNPIQTKSIPIKGENECVFQVLP-----GRLYSVTVTTK 803  
Db 294 KNNWEEQSGNFP-----EESPMRQDTIGKEKLPHEFTEETPEIPSGNISGWPDPFN 345  
Qy 804 SQQYEAHQ-----GNRTIPEP-----VKDLTLNRSTEDLHVTVSGANGVDQYEQIQLP 855  
Db 346 SSDYETTSQPYWDSASAPESEDFVSVLPMYENNSTLSEKSTSGSFPFPQML 405  
Qy 856 NDMKVFP-----PHLVNTATEYRTSL-----TPGROYKILVLTISGDV 895  
Db 406 TWLPKPPTADGPH-IHIEREENFETYLWDEAEHFAELKEPGK-YKLSVTFS--- 460  
Qy 896 QCSAFIEGTVPFAVK-NIHISPGATDSLTVNTPGGGDVDSYTVSAFRHSQKVDQTI 954  
Db 461 -SSGCETKRSQSAKSLFYISPSG-----EW-----IEELTEK----- 493  
Qy 955 PKHVFETHFRLAGEQYQIMIASVSGSLKNQINNVGRTVPASVQGVADNAYSYSILV 1014  
Db 494 PQHVSVHLSTTA-----LMSWTSQENYNTI-----VSVSLTC 530  
Qy 1015 SWQKAAGVAERYDILLTENGILLRNTSEPTATKQKPEDLTGKYYKIQLITVSGGLFS 1074  
Db 531 QKQKE---SQLEKQYCTQ-----VNSSKPI-----IENLVGAQYQVVIYLRKGPLIG 576  
Qy 1075 KEAQ-TEGRTVPAAVTDLIRITENSTRHLSFRWTASE-GELSWY--NIFLYNPDGNLQERA 1130  
Db 577 PPSDPVTAIVPTGIKDLMLPLGPTAVVLSWTRPYLGVFRKYVEMFNPATMSEWT 636  
Qy 1131 QVDPLVQSFQNLLOGRMYKMWIVTHSGELSNESFIFGRTVPASVSHLGRSNRTTDSL 1190  
Db 637 TYYEIAATVS-----LTASVRIAN-----LLPA----- 659  
Qy 1191 WFNWSPASGDFDYELIILYNPNTGKKNKDKLTWFRQGLVPCRKYVLVWVTHSGDLS 1250  
Db 660 WY-----YNFRVT-MVTWGDPELS-----CC 679  
Qy 1251 NKVTAESRTAPSPSLMSFADIANSTLAIWKGPDPDWTYDNDFE-LOWLPDALTYFNPY 1309  
Db 680 DSSTISFITAPVAPETIS-VEYFNLSLYISWTYGGDDTTDLSHSRMLHW----- 727  
Qy 1310 NNRKSEGRVGLRPGRSYQFNVTVSGDSWKTYSKPIFGSVRT-----KPKIQLNLCR 1364  
Db 728 -----VWAEKKKIKKSVTRNVMTAILSLPGDIYNL----- 759  
Qy 1365 PQNSTAIACSWIPDPDFDGSIECRKMDTQVEFSRKLKEKSLNINMLVPHKRYLVS 1424  
Db 760 ----SVTACT-----ERGSNTSMRLVKL----- 779

Qy 1425 IKVQSAGTSEVEDSTITWIDRPPPPPHIRVNEKDVLSKSSINFVNCWSPDTNGA 1484  
Db 780 -----EPAPPKSLFVANKTQT-----SVTLWLWEE--GV 806  
Qy 1485 VKYFTVVVRADGSDLEKPEOQHPLPSVLEYRHNASIRVYQTNYPASKCAENPNNSKSF 1544  
Db 807 ADFFEVCQGVSSQKTKLQEPVAVSS-----HVVTI----- 838  
Qy 1545 NIKLGAEMESLGGKRDPTQOKFCGDKPLKPHYATIRASTRAFTQLDFDEKFTKPLYSDTF 1604  
Db 839 -----SSLLPATAYNCVTSFSH-----DSPSV-PTF 864  
Qy 1605 PSL-PITTESEPLEGATEGVS-AGLFLIGMLVAVALLICROK-VSHGRERPSARL----- 1657  
Db 865 IAVSTMVTEMNPNVIVISLAILSTLLIGLL--LVTIILRKKGLQWARECAGCTFVNFA 922  
Qy 1658 SIRDRLPSVHLNLGQGNRKTSCPIKINOPEGHFMLQADSNYLLSKEYBELKDVGRNQ 1717  
Db 923 SLEBDGKLPYNWSKNGIKKRLNPNVQDDPDAYIKWAKDSYKFSLOFELKILGLDI 982  
Qy 1718 SCDIALLPENRGNRYNNILPYDATRVKLSNVDDPCSDYINASYIPGNFRREYIVTQG 1777  
Db 983 PHFAADLPNCKNRYTNILPYDFSRLVRLVSMNEEGADYINANYIPGYNPQEVIAVQ 1042  
Qy 1778 PLPTKDDDFWQWQEQNVNHNIVMVTQCVKGRVKCDHYWADQDSLYYDGLLOMLSESV 1837  
Db 1043 PLPETRNDFWKVLQOKSQIIVMLTQCNKERVKCDHYWPTTEPIAYGDTIVEMISSE 1102  
Qy 1838 LPEWTIREFKICEQEQLDAHLRHFHYVTWPDHGV--ETQSLIQPVRTVRDYNRSP 1895  
Db 1103 QDDWACHRFI---NYADEMQDVNHFNTAWPDHGVPTANAAESILQFVHMVROQATKS- 1159  
Qy 1896 GAGTVVHCAGVGRGTGTFIALDRLQQLDSKSDVIYGAHVHDLRLHRVHMVQTECQVY 1955  
Db 1159 -KGMIIHCSAGVGRGTGTFIALDRLQHLRHFHFDVILGLVSEMRYSMSVQTEEQYIF 1217  
Qy 1956 LHQCRDVLRLARKURSEQENPLFFIYENVN 1985  
Db 1218 IHQCVQLMMWKKK---QQFCISDIYVENVS 1244

## RESULT 14

US-09-949-016-8052  
; Sequence 8052, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8052  
; LENGTH: 1246  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-8052

Query Match 8.7%; Score 913.5; DB 2; Length 1246;  
Best Local Similarity 24.0%; Pred. No. 2.1e-49;  
Matches 367; Conservative 199; Mismatches 483; Indels 481; Gaps 63;  
Qy 571 LFTNWTQAGDVEFYQVLLIHENVIKNES-----ISSETSRY----- 608

Db 81 LFKNAT-----AFHVTVDQNNIVVSLASDVSPASVYVVKITGBSKNYFFBFEFNS 134  
Qy 609 -----SFHSLKSGSLSVVVTTVSGGI-----SSRQVVVEGRTVPSSVSGVTVNSG 655  
Db 135 TLPPVPVPKASH-----GLYIITLVVNGVNVTKPSRSITV--LTKPLPVTSVIYDYK 188  
Qy 656 RNDYLSVSWLAPGDVDNY--EVLISH-DGKVQSLVI-----AKSVRECSFSSLTGRL 707  
Db 189 PSPETGVLFETHPEKYNVFRVNIISYWEKDFRMLYKDFKGTIV---FNHVLQMC 244  
Qy 708 YT-----VTITRSGKYENHSPSQBERTVPDKVQGVSVNSARSYDLRV--- 750  
Db 245 YSNITFQLVSEATFNKSLVEYSG--VSHPEKQHTAPYPPQNISV-----RIVNL 293  
Qy 751 ---SWHATGDFHVEVITIKKNFIQTKSPKSENECVFQVLP-----GRLYSVTVTK 803  
Db 294 KNNWEEOSGNFP-----EESFMRSQDTIGKEKLFHTEETPEIPSGNISGWPDFN 345  
Qy 804 SQQYEANEQ-----NGRRTIPEP-----VKDLTLRNRSTEDLHVTWGGANGDVQYEIQLLF 855  
Db 346 SSDYETTSQPYWMDSASAPESEDEFVSVLPMYEYNNSTLSETEKSTSGSFSPVQML 405  
Qy 856 NDMKVFP-----PHLVNTATEYFTSL-----TPGRQYKILVLTISGDV 895  
Db 406 TWLPPKPTAFDGFH-IHIEREENFTEYLMVDEBAHEFVABELKEPGK-YKLSVTTFS--- 460  
Qy 896 QOSAFIEGFTVPSAVK-NIHISPNGATDSLTVNTPCGGDVDSYTVSAFRHSQKVDQTI 954  
Db 461 -SSGCETRKSQSAKSUSFYISPSG-----EW-----IEELTEK----- 493  
Qy 955 PKHVFEHTFHRLEAGEQOIMIASVSGSLKNOINVVGRTVPASVOGIADNAYSYSILV 1014  
Db 494 PQHVSVHLSSTTA-----LMWTSQENYNTI-----VSVSLTC 530  
Qy 1015 SWQAAQVAERYDILLTENGILLRNTSEBATTQKHKPEDTPGKQYKIQLTVSGGLFS 1074  
Db 531 QKQKE---SORLEQYCTQ-----VNSKPI-----IENLVGAQYQVVIYLRKGLIG 576  
Qy 1075 KEAQ-TEGRTPAAVTLRITENSTRHLSFWTASE-CELSWY--NIFLYNPDGNLQERA 1130  
Db 577 PPSDPVTFATVPTGIKODMLPLGFTAVLSWTRPYLGFRKYVVFYFNPAATMTSEWT 636  
Qy 1131 QVDPVLVSFSPQNLQGRMYKMWIVTHSGELSNESFIFGRTPVPASVSHLRGSRNRTDLSL 1190  
Db 637 TYEIAATVS-----LTASVRAN-----LLPA----- 659  
Qy 1191 WFNWSPASGDPDFVELILYNPNGTCKENWKDKDLEWRFOGLVPRKYVLVWVTHSGDLS 1250  
Db 660 WY-----YNRVPT-MVTWGDPELS-----CC 679  
Qy 1251 NKVTAERTAPSPSLMSFADIANTSLAITWKGPDMDTYNDPE-LQWLPRDALTVFNPY 1309  
Db 680 DSSITSIFITAPVAPEITS-VEYFNSLYISWTYGGDDTDLJSHSRMLHWM----- 727  
Qy 1310 NNRKSEGRVYGLRPGRSQFNVTVDGSDSKTSKPIFGSVRT-----KPKDKIONLHCR 1364  
Db 728 -----VVAEGKKIKKSVTRNVMTAILSLPPGDIYNL--- 759  
Qy 1365 PNSTAIACSWIPPDSPDFDGYSEICRKNVDTOEVBFSRKLKESLNLMLVPHKRYLVS 1424  
Db 760 ---SVTACT-----ERGSNTSMLRLVKL----- 779  
Qy 1425 IKVQASCTSEVDESDITIMIDRPPPPPHIRVNEKVDLSKSSINTVNCWPSDTNGA 1484  
Db 780 -----EPAPPKSLFANVKTQT-----SVTLWVEE--GV 806  
Qy 1485 VKYFTVVVREADGSDLEKPEOQHPLPSYLEVRHNASIRVYQTNVFSKCAENPNNSKSF 1544  
Db 807 ADPEVEVCQVGSQKTKLQEPVAVSS-----HVVTI----- 838  
Qy 1545 NIKLGAEMESLGGKRDPTQOKFCDDGLPKPHPTAYISIRAFQLPDEDKFTKPLYSDTF 1604  
Db 839 -----SLLPATAYNCVTSFSH-----DSPSV-PTF 864

Qy 1605 FSL-PITTESEPLFGAIEGVS-AGLFLIGHLVAVALLICROK-VSHGRBRPSARL---- 1657  
Db 865 IAVSTWTEWNPVNVVLSJLAILSTLLIGLT--LVTLIIIRKHLQWARECGAGTFVNF 922  
Qy 1658 SIRDRPLSVHLNLGQKGNKRTSCPIKINOFEGHFMKLQADSNYLLSKEYEELKXDVGRNQ 1717  
Db 923 SLERDGLPYNWSKNGLKKRKLNPVQLDDFDAYIKDMAKSDYKFSLQFEELKLGID 982  
Qy 1718 SCDIALLPENRGKRYNNILPYDATRVKLSNVDDDDPCSDYINASYIFGNPRRYIYVTOG 1777  
Db 983 PHFAADPLNCRKNRYNILPYDFSRVRLVSMNEEGADYINANYIFGYNPSQFYIATQG 1042  
Qy 1778 PLCTKDDFKWQWQENVHNVMTQCVKEGRVKCDHYWADODSLYVGLLILQMLSES 1837  
Db 1043 PLPEIRNDFWMLVQKQKQIIVMLTQCNERRKVKCDHYWPTTEPIAYGIDITVMISEEE 1102  
Qy 1838 LPEWTIRREFKICGEBEQDAHRLIRHFHYTVMPDHGVP--ETTOSLIQFVTRTVRDIYNRSP 1895  
Db 1103 QDDWACRHFRI---NYADEMQDVMHFNYTAMPDGHVPTANAEISILQFVHMVRQOATKS- 1158  
Qy 1896 GAGTIVVHCAGVGRITGTFIALDRILQOLDSKDSVDIYGAVHDLRLHRVHMVQTECOYV 1955  
Db 1159 -KGMIIHCSAGVGRITGTFIALDRILQHIRDHEFDVILGLVSEMRSYRMSMVQTEEQYIF 1217  
Qy 1956 LHOCVRDLRLARKLRSQENPLPPIYENVN 1985  
Db 1218 IHOCVQLMWMKKK---QQFCISDVIYENVN 1244

## RESULT 15

US-09-949-016-8053  
; Sequence 8053, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8053  
; LENGTH: 1246  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-8053

Query Match 8.7%; Score 913.5; DB 2; Length 1246;

Best Local Similarity 24.0%; Pred No. 2.1e-49;

Matches 367; Conservative 199; Mismatches 483; Indels 481; Gaps 63;

Qy 571 LFTNWTQAOQDVEFYQVLLIHENVVKNES-----ISSTSRY----- 608  
Db 81 LFKNAT-----APHVTVQDDNNIVVSLASDVSPASVYVVKITGBSKNYFFBFEFNS 134  
Qy 609 -----SFHSLKSGSLSVVVTTVSGGI---SSRQVVVEGRTVPSSVSGVTVNSG 655  
Db 135 TLPPVPVPKASH-----GLYIITLVVNGVNVTKPSRSITV--LTKPLPVTSVIYDYK 188  
Qy 656 RNDYLSVSWLAPGDVDNY--EVLISH-DGKVQSLVI-----AKSVRECSFSSLTGRL 707  
Db 189 PSPETGVLFETHPEKYNVFRVNIISYWEKDFRMLYKDFKGTIV---FNHVLQMC 244  
Qy 708 YT-----VTITRSGKYENHSPSQBERTVPDKVQGVSVNSARSYDLRV--- 750



```
Db 245 YSNITFQLVSEATFNKSTLIVEYSG--VSHEPKQHRTAPYPPQNISV-----RIVNL 293
Qy 751 ---SWHATGDDHVEYIKNNFIQTKSPKSENECVFQVLP-----GRYSVTVTTK 803
Db 294 NKNWEEQSGNFP-----EESFMRSDTIGKEKLFHTEETPEIPSGNISSGWPDEN 345
Qy 804 SQQYANEQ---GNRTIPEP---VKDLTLRNSTEDLHVWTSAGNDVDQYBIQLLF 855
Db 346 SSDYETTQPYWDSASAPSESEDFVSLPMEYENNSTLSEKSTSGSFSPFQVMIL 405
Qy 856 NDMKVPPP-----PHLVNATREYFTSL-----TEGROYKILVLTISGDV 895
Db 406 TWLPKPTAPDGFH-IHIERENFTEYLVWDEBAHEFVAELKEPGK-YKLSVTTFSS--- 460
Qy 896 QQSARIEGFTVPSAVK-NIHISPGATSLTNWTPGGDVDSYTVSAFRHSQKVDSTI 954
Db 461 -SSGSCETRKQSASKLSFYISPSG-----EW-----IEBELTEK----- 493
Qy 955 PKHVFHTFHRLEAGEQIIMIASVSGSLKNQINVVGRTPVPSVQGVADNAYSYSILV 1014
Db 494 POHVSVHVSSTTA-----LMSWTSQENTNSTI-----VSVSLTC 530
Qy 1015 SWQKAAGVAERYDILLTENGILLRNTSEPAATTKQHKFEDLTPGKKYKIQLTVSGGLFS 1074
Db 531 QKQKE---SORLEKQYCTQ-----VNSKPI-----IENLVPGAQYQWVIYLRKGLIG 576
Qy 1075 KEAQ-TEGRTPAAVTDLRITENSTRHLSFRWTASE-GELSWY--NIFLYNPDGNLQERA 1130
Db 577 PPSDPTFAIVPTGIKDLMYPLGPTAVVLSWTRPYLGVFRKYVVMFYFNPATMTSEWT 636
Qy 1131 QVDPPLVQSFSPQNLQGRMYKMWIVTHSGELSNESFIFGRTPVPSVSHLGRSNRTTDSL 1190
Db 637 TYIEIAATVS-----LTASVRIAN-----LLPA----- 659
Qy 1191 WFNWSPASGDFDYELIILYNPNTKKENWKDLETWRFGQLVPGKRYVLVWVTHSGDLS 1250
Db 660 WY-----YNFRVT-WVTWGPDELS-----CC 679
Qy 1251 NKVTABSTAPSPSLMFADIANSTLAIWKGPDPDWTYNDNFE-LQWLPRDALTVFNPY 1309
Db 680 DSSTISFITAPVAPEITS-VEYFNLLYISWTYGDYDITDLSHRMLHWN----- 727
Qy 1310 NNRKSEGRIVYGLRGRSVQFNWKTVSGDSWKYSPKPIFGSVRT-----KDDKIONLHCR 1364
Db 728 -----VWAEGRKKIKKSVTRNVMTAILSLPPGDIYNL--- 759
Qy 1365 PQNSTAIACSWIPPPDSDFGYSIECKRMDTQEVFSPKRLEKEKSLNIMMLVPHKRYLVS 1424
Db 760 ----SVTACT-----ERGSNTSMRLVKL----- 779
Qy 1425 IKVQAGMTSEVEDSTITMIDRPPPPPHIRVNEKQVLIKSXSINFVNCWSFSDTNGA 1484
Db 780 -----EPAPPKSLFAVNKTQT-----SVTLWVBE--GV 806
Qy 1485 VKYFTVVVREADGSDLPQHQHPLPSPVLEYRHNASIRVQTNYPASKCAENPNSKSF 1544
Db 807 ADFFEVCQVQSGSOKTLOEPVAVSS-----HVVTI----- 838
Qy 1545 NIKLGAEMESLGGKRDPTQKFCDGKPLKPHATYRISIRAFQTQLPDEDLKEFTKPLYSDTF 1604
Db 839 -----SSLLPATAYNCVTSFSH-----DSPSV-PTF 864
Qy 1605 FSL-PITTESEPLFAGBVS-AGLFLIGMLVAVVALLICRQK-VSHGRERPSARL----- 1657
Db 865 IAVSTWVTIEMPNVVVISVAILSTLLIGLL--LVTLILRKHLQMAECGAGTFVNFA 922
Qy 1658 STRRDRPLSVHLNLOKGNKRTSCPIKINQEGHFMKLQASNYLLSKEYBELKDVGRNQ 1717
Db 923 SLERDGLPYNWSKNGKLKRLTNPVQLDDFDAYIKDMAKDSYKFSLQFELKILGDI 982
Qy 1718 SCDIALLPNRCKRNYNITLPYDATRVKLSNVDDPCSDYINASYIPGNFRREYIVTQG 1777
Db 983 PHFAADLPLNRCKRNYNITLPYDFSRVRLVSNWEEGADYINANYIPGNSPQEIATQG.1042
```

Search completed: March 10, 2006, 18:39:37

Job time : 62.7221 secs

This Page Blank (uspto)

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2006, 18:59:00 ; Search time 196.859 Seconds  
(without alignments)  
4238.599 Million cell updates/sec

Title: US-10-633-742-2

Perfect score: 10483

Sequence: 1 MLSHGAGLALWITLSLLQTG.....FPIYENVPEYHRDPVYSRH 1997

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : 1: Pubcgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pap.\*

2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pap.\*

3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pap.\*

4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pap.\*

5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pap.\*

6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10483	100.0	1997	4	US-10-634-027-2
2	10480	99.9	1997	5	US-10-756-149-5168
3	10460	99.8	1997	3	US-09-909-567B-54
4	10460	99.8	1997	4	US-10-408-765A-2135
5	10460	99.8	1997	5	US-10-497-692-4
6	7614	72.6	1450	5	US-10-497-692-14
7	1830	17.5	336	4	US-10-634-027-4
8	1691	16.1	319	4	US-10-634-027-7
9	1690	16.1	312	4	US-10-634-027-6
10	1556.5	14.8	1647	6	US-11-097-143-2076
11	1485.5	14.2	1767	4	US-10-087-684-40
12	1485.5	14.2	1767	4	US-10-218-779-40
13	1485.5	14.2	1767	6	US-11-097-143-14970
14	1475.5	14.1	1767	4	US-10-087-684-41
15	1475.5	14.1	1767	4	US-10-218-779-41
16	1433.5	13.7	1447	5	US-10-497-692-13
17	1374	13.1	310	3	US-09-788-626-15
18	1247	11.9	1337	4	US-10-390-501-2
19	1243.5	11.9	1238	4	US-10-366-547-47
20	1242	11.8	1337	4	US-10-366-547-42
21	1242	11.8	1337	4	US-10-723-606-2
22	1221.5	11.7	1216	4	US-10-366-547-49
23	1206	11.5	1337	4	US-10-366-547-44
24	1204	11.5	1705	5	US-10-029-345A-37
25	1198	11.4	1711	5	US-10-029-345A-38
26	1107	10.6	2291	4	US-09-822-871-2
27	1107	10.6	2291	4	US-10-673-885-2

28	1106	10.6	2281	4	US-10-087-684-6	Sequence 6, Appli
29	1106	10.6	2281	4	US-10-218-779-6	Sequence 6, Appli
30	1100.5	10.5	2300	4	US-10-087-684-10	Sequence 10, Appli
31	1100.5	10.5	2300	4	US-10-218-779-10	Sequence 2, Appli
32	1099.5	10.5	2299	4	US-10-466-759-2	Sequence 4, Appli
33	1097	10.5	2301	3	US-09-822-871-4	Sequence 4, Appli
34	1097	10.5	2301	4	US-10-673-885-4	Sequence 37, Appli
35	1097	10.5	2302	4	US-10-087-684-37	Sequence 37, Appli
36	1097	10.5	2302	4	US-10-218-779-37	Sequence 5508, Ap
37	939	9.0	1367	4	US-10-369-493-5509	Sequence 5509, Ap
38	939	9.0	1367	4	US-10-331-496A-79	Sequence 79, Appli
39	913.5	8.7	1188	4	US-10-789-241-34	Sequence 34, Appli
40	913.5	8.7	1188	4	US-10-029-345A-27	Sequence 27, Appli
41	913.5	8.7	1188	5	US-10-723-606-3	Sequence 3, Appli
42	889.5	8.5	341	4	US-10-245-539-4	Sequence 2, Appli
43	886	8.5	1118	4	US-10-245-539-2	Sequence 2, Appli
44	886	8.5	1118	4	US-10-245-539-8	Sequence 8, Appli
45	886	8.5	1118	4	US-10-245-539-8	Sequence 8, Appli

#### ALIGNMENTS

##### RESULT 1

US-10-634-027-2

; Sequence 2, Application US/10634027

; Publication No. US20040077065A1

; GENERAL INFORMATION:

; APPLICANT: Procter & Gamble Company

; APPLICANT: Evdokimov, Artem G

; TITLE OF INVENTION: Three Dimensional Coordinates of HPTbeta

; FILE REFERENCE: 9045M2

; CURRENT APPLICATION NUMBER: US/10/634, 027

; PRIOR FILING DATE: 2003-08-04

; PRIOR APPLICATION NUMBER: US 60/413,547

; PRIOR FILING DATE: 2002-09-25

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: Patent in version 3.2

; SEQ ID NO 2

; LENGTH: 1997

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-634-027-2

Query Match 100.0%; Score 10483; DB 4; Length 1997;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1997; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLSHGAGLALWITLSLLQTGLAEPERCNFTLAESKASSHSVSQWRILGSPCNFSLIYSS 60

Db 1 MLSHGAGLALWITLSLLQTGLAEPERCNFTLAESKASSHSVSQWRILGSPCNFSLIYSS 60

Qy 61 DTGGAALCPTRFDNTTYGNCLOAGTIYFNKILSLDEERTVLTQDLPAPFVSK 120

Db 61 DTGGAALCPTRFDNTTYGNCLOAGTIYFNKILSLDEERTVLTQDLPAPFVSK 120

Qy 121 EKTSTGLHWMTPTSSGKVTSEYVQLFDENNQIQGVQIOESTSWNEYTFNLTAGSKYN 180

Db 121 EKTSTGLHWMTPTSSGKVTSEYVQLFDENNQIQGVQIOESTSWNEYTFNLTAGSKYN 180

Qy 181 IATVAVGGKRSRPSVTNGSTVPSVKDIGISTKANSLIISWHSNGSNVRYRLMLMDKG 240

Db 181 IATVAVGGKRSRPSVTNGSTVPSVKDIGISTKANSLIISWHSNGSNVRYRLMLMDKG 240

Qy 241 ILVHGGVVDKHTSYAFHGLSPCYLYNLVTMTAAQLQNYRWKLVRTAPNEVSNLKYTND 300

Db 241 ILVHGGVVDKHTSYAFHGLSPCYLYNLVTMTAAQLQNYRWKLVRTAPNEVSNLKYTND 300

Qy 301 GSLSLTKVQWQRPNGVDSYNTLSHKGTIKESRLAPWITETHFKELVPGRLYQVTVSC 360

Db 301 GSLSLTKVQWQRPNGVDSYNTLSHKGTIKESRLAPWITETHFKELVPGRLYQVTVSC 360

```
Qy 361 VSGELSAQKMAVGRTPDPKQVANLEANNNGRMRSILVSVKSPAGDWEQYRILLFNDSSVLL 420
Db 361 VSGELSAQKMAVGRTPDPKQVANLEANNNGRMRSILVSVKSPAGDWEQYRILLFNDSSVLL 420
Qy 421 NITVGKEETOYVMDDTGLVPGROQVEVEVIVESGNLKNSERCQGRTPVPLAVQLRVKANE 480
Db 421 NITVGKEETOYVMDDTGLVPGROQVEVEVIVESGNLKNSERCQGRTPVPLAVQLRVKANE 480
Qy 481 TSLSIMMQTPVAEWEKYIISLADRDLLLIHKLSKDAKEFTFDLPGRKYMATVTSISG 540
Db 481 TSLSIMMQTPVAEWEKYIISLADRDLLLIHKLSKDAKEFTFDLPGRKYMATVTSISG 540
Qy 541 DLKNSSSVKGRTPVPAQVTDLHVANQGMTSSLFTNTWTAQGDVEFYQVLLIHENVVIKNES 600
Db 541 DLKNSSSVKGRTPVPAQVTDLHVANQGMTSSLFTNTWTAQGDVEFYQVLLIHENVVIKNES 600
Qy 601 ISSETSYSHLSKSGSLYSVVVTVTSSGGISSRQVVVEGRTVPSSVSGVTVNNSGRNDYL 660
Db 601 ISSETSYSHLSKSGSLYSVVVTVTSSGGISSRQVVVEGRTVPSSVSGVTVNNSGRNDYL 660
Qy 661 SVSWLVAPGDVDNVEVTLSDHGKVVQSLVIAKSVRECSFSSLTPGRLYTVTITTRSGKYE 720
Db 661 SVSWLVAPGDVDNVEVTLSDHGKVVQSLVIAKSVRECSFSSLTPGRLYTVTITTRSGKYE 720
Qy 721 NHPFSQERTVPDKVQGVSVNSARSYLRSVSWHATGDFDHYEYTIKNKNFIQTKSIPK 780
Db 721 NHPFSQERTVPDKVQGVSVNSARSYLRSVSWHATGDFDHYEYTIKNKNFIQTKSIPK 780
Qy 781 SENECVQVLPGRLYSVTVTKSGQYEAEOQNGRTIPEPVKDLTLNRSTEDLHVTVS 840
Db 781 SENECVQVLPGRLYSVTVTKSGQYEAEOQNGRTIPEPVKDLTLNRSTEDLHVTVS 840
Qy 841 GANGDVQYEQIQLLFNDKMFPPPHLVNTATEYRFTSLTPGRQVKILVLTISGDVQOSAF 900
Db 841 GANGDVQYEQIQLLFNDKMFPPPHLVNTATEYRFTSLTPGRQVKILVLTISGDVQOSAF 900
Qy 901 IEGFTVPSAVKNTHISPNGATDSLTVNWTPEGGDVDSYTVSAFRHSQKVDSTQIPKHVFE 960
Db 901 IEGFTVPSAVKNTHISPNGATDSLTVNWTPEGGDVDSYTVSAFRHSQKVDSTQIPKHVFE 960
Qy 961 HTFHRLBAGQYQIMIASVSGSLKNQINNVGRTPVPSVQGVIAADNAYSLSLIVSQKAA 1020
Db 961 HTFHRLBAGQYQIMIASVSGSLKNQINNVGRTPVPSVQGVIAADNAYSLSLIVSQKAA 1020
Qy 1021 GVAERYDILLTENGILLRNTSEBATTQKHFEDELTPGKKYKQILTVSGGLSKQAQTE 1080
Db 1021 GVAERYDILLTENGILLRNTSEBATTQKHFEDELTPGKKYKQILTVSGGLSKQAQTE 1080
Qy 1081 GRTVPAAVTDLRITENSTRHLSFRWTASEGELSWYNIFLYNPDGNLQERAQVDPVQSFS 1140
Db 1081 GRTVPAAVTDLRITENSTRHLSFRWTASEGELSWYNIFLYNPDGNLQERAQVDPVQSFS 1140
Qy 1141 FQNLQGRMYKMWIVTHSGELSNESFTFGRTVPAASVSHLRGSNRNTDLSLWFNWPASGD 1200
Db 1141 FQNLQGRMYKMWIVTHSGELSNESFTFGRTVPAASVSHLRGSNRNTDLSLWFNWPASGD 1200
Qy 1201 FDFVELLYNPNGTKEKNKDKDLTEWRFOGLVPGRYKVLVWVTHSGDLNKNKYTAESRTA 1260
Db 1201 FDFVELLYNPNGTKEKNKDKDLTEWRFOGLVPGRYKVLVWVTHSGDLNKNKYTAESRTA 1260
Qy 1261 PSPPLSMFADIANTSIAITWKGPDPDWTYDNDFELQWLPRDALTYVFNPNYNNRSEGRIVY 1320
Db 1261 PSPPLSMFADIANTSIAITWKGPDPDWTYDNDFELQWLPRDALTYVFNPNYNNRSEGRIVY 1320
Qy 1321 GLRPGRSYQNVKTVSGDSWKTSYKPIFGSVRTKPKDIQNHLHCRPQNSTAIACSWIPPDSS 1380
Db 1321 GLRPGRSYQNVKTVSGDSWKTSYKPIFGSVRTKPKDIQNHLHCRPQNSTAIACSWIPPDSS 1380
Qy 1381 DFDGYSIECKRMDTOEVEFSPKLEKEKSLANIMWLVPHKYLYSVIKVQSAGMTSEVVEDS 1440
Db 1381 DFDGYSIECKRMDTOEVEFSPKLEKEKSLANIMWLVPHKYLYSVIKVQSAGMTSEVVEDS 1440
Qy 1441 TITMIDRPPPPPHIRVNEKDVLSISKSSINFVNCNCSWFDSTNGAVKYFTVVVREADGSDE 1500
```

```
Db 1441 TITMIDRPPPPPHIRVNEKDVLSISKSSINFVNCNCSWFDSTNGAVKYFTVVVREADGSDE 1500
Qy 1501 LKPEQOHPLESLYLEYRHNASIRVYOTNYFASKCAENPNSKSFNIKLGAEMESLGKRD 1560
Db 1501 LKPEQOHPLESLYLEYRHNASIRVYOTNYFASKCAENPNSKSFNIKLGAEMESLGKRD 1560
Qy 1561 PTQOKFCGDKLPDHTAYRISIRAFQTOLFDEDLKEFTKPLYSDFFSPLPITTESEPLFGAI 1620
Db 1561 PTQOKFCGDKLPDHTAYRISIRAFQTOLFDEDLKEFTKPLYSDFFSPLPITTESEPLFGAI 1620
Qy 1621 EGVSAGLFLIGMLVAVVALLICRQKVSCHGRERPSARLSIRDRPLSVHLNLGQGNRKTS 1680
Db 1621 EGVSAGLFLIGMLVAVVALLICRQKVSCHGRERPSARLSIRDRPLSVHLNLGQGNRKTS 1680
Qy 1681 CPKINQFEGHFMKLOADSNYLLSKEYEELKDVGRNQSCDIALLPENRGKRYNNILPYD 1740
Db 1681 CPKINQFEGHFMKLOADSNYLLSKEYEELKDVGRNQSCDIALLPENRGKRYNNILPYD 1740
Qy 1741 ATRVKLSNVDDDDPCSDYINASYIPGNFRREYIVTQGPLPGTKDDDFWQVWQVNHVIM 1800
Db 1741 ATRVKLSNVDDDDPCSDYINASYIPGNFRREYIVTQGPLPGTKDDDFWQVWQVNHVIM 1800
Qy 1801 VTQCVKGRVKCDHYWPADQDSLYYGDLLQMLSESVLPPEWTIREFKICGEEQLDAHRLI 1860
Db 1801 VTQCVKGRVKCDHYWPADQDSLYYGDLLQMLSESVLPPEWTIREFKICGEEQLDAHRLI 1860
Qy 1861 RHPHYTVWPDHGVPEPTQSLIQFVTRVDYINRSPGAGPTVHCHSAGVGTCTFIALDRI 1920
Db 1861 RHPHYTVWPDHGVPEPTQSLIQFVTRVDYINRSPGAGPTVHCHSAGVGTCTFIALDRI 1920
Qy 1921 LQQLDSKDSVDIYGAVHDLRLHRVHMVQTECOVYVYLHQCVRDVLRLARKLRSEQENPLFPI 1980
Db 1921 LQQLDSKDSVDIYGAVHDLRLHRVHMVQTECOVYVYLHQCVRDVLRLARKLRSEQENPLFPI 1980
Qy 1981 YENNVPEYHRDPVYSRH 1997
Db 1981 YENNVPEYHRDPVYSRH 1997

RESULT 2
US-10-756-149-5168
; Sequence 5168, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5918
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5168
; LENGTH: 1997
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-756-149-5168

Query Match 99.9%; Score 10480; DB 5; Length 1997;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1996; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLSHGAGLAIWITLSLLOTLGLABPERCNETFLAESSKASSHSVSIQWRILGSPCNFSLIYSS 60
Db 1 MLSHGAGLAIWITLSLLOTLGLABPERCNETFLAESSKASSHSVSIQWRILGSPCNFSLIYSS 60
Qy 61 DTUGAALCPTFRIDNTTYGNCNLOAQGTIYNPKIISLDEERTVVLQTDPLPPARFQVSK 120
Db 61 DTUGAALCPTFRIDNTTYGNCNLOAQGTIYNPKIISLDEERTVVLQTDPLPPARFQVSK 120
Qy 121 EKTSTGLHWWTPSSSGKVTSYEVQLFDENNQKIQGVQIQESTSWNEYTFNLTAGSKYN 180
```

121 EKTSTGLHWWTPSSSKVTSYEQVLEFENNQKIQGVQIQESTSWNEYTFNLTAGSKYN 180  
181 IAITAVSGGRKRSFVYTNSTGSPVVKDIGISTKANSLLISWSHGSGNVERYRLMLMDKG 240  
181 IAITAVSGGRKRSFVYTNSTGSPVVKDIGISTKANSLLISWSHGSGNVERYRLMLMDKG 240  
241 ILVHGGVVDGHTSYAFHGLSPGYLYNLTYMTAEAGLQNYRWKLVRTAPMEVSNLKVTDND 300  
241 ILVHGGVVDGHTSYAFHGLSPGYLYNLTYMTAEAGLQNYRWKLVRTAPMEVSNLKVTDND 300  
301 GSLSLTKVKKQRPNGNDVSNITLSHGKGTIKESRVLAPMITETHFKELVPGRLYQVTVSC 360  
301 GSLSLTKVKKQRPNGNDVSNITLSHGKGTIKESRVLAPMITETHFKELVPGRLYQVTVSC 360  
361 VSGELSAQKMAVGRTPPDKVANLEANNNGMRSLVSWSPAGDQWQYRILLFNDSVLL 420  
421 NITVKGEEQVMDDTGLVPGROYEVEVIVESGNLKNSERCQGRTPVLAQLVRKHANE 480  
421 NITVKGEEQVMDDTGLVPGROYEVEVIVESGNLKNSERCQGRTPVLAQLVRKHANE 480  
481 TSLSIMMOTPVAEWEKIIISLADRDLLIIHKSLSKDAKEFTFDLPGRKYNATVTSISG 540  
481 TSLSIMMOTPVAEWEKIIISLADRDLLIIHKSLSKDAKEFTFDLPGRKYNATVTSISG 540  
541 DLKNSSSVKGRTPVPAQVTDLFVANQGMTSSLFTNTWTOAQGDVEFYQVLLIHENVIKNES 600  
541 DLKNSSSVKGRTPVPAQVTDLFVANQGMTSSLFTNTWTOAQGDVEFYQVLLIHENVIKNES 600  
601 ISSETSYSHLSKSGSLYSVWVTTVSGGSISSROVVVEGRTVPSSVSGVTNNSGRNDYL 660  
601 ISSETSYSHLSKSGSLYSVWVTTVSGGSISSROVVVEGRTVPSSVSGVTNNSGRNDYL 660  
661 SVSWLAPGDVNDYEVTLSDHGKVVQSLVIAKSVRECSFSLTPGRUYTYTITRSKYE 720  
661 SVSWLAPGDVNDYEVTLSDHGKVVQSLVIAKSVRECSFSLTPGRUYTYTITRSKYE 720  
721 NHSFQSBQRTVPDKVQSVNSARSYLRSVSWHATGDFDHYEYTIKKNFNFIQTSIPK 780  
721 NHSFQSBQRTVPDKVQSVNSARSYLRSVSWHATGDFDHYEYTIKKNFNFIQTSIPK 780  
781 SENECPVQVLPGRLYSVVTTVSGQYEAQNGNRTIPBPVKDILNRSSTEDLHVTWS 840  
781 SENECPVQVLPGRLYSVVTTVSGQYEAQNGNRTIPBPVKDILNRSSTEDLHVTWS 840  
841 GANCVDVQYIOLLFNDKMFPPPHLVNTATEYRFTSLTPGRQYKILVLTISGDVQOSAF 900  
841 GANCVDVQYIOLLFNDKMFPPPHLVNTATEYRFTSLTPGRQYKILVLTISGDVQOSAF 900  
901 IEGFTVPSAVKNIIHISPNAGTDSLTVNWTGGGDVDSYTVSAFRHSOKVDSQTIPIKHVFE 960  
901 IEGFTVPSAVKNIIHISPNAGTDSLTVNWTGGGDVDSYTVSAFRHSOKVDSQTIPIKHVFE 960  
961 HTFHRLEAGQYQIMIASVSGSLKNQINNVGRTVPASVQGVIAADNAYSYSLLIVSWQKAA 1020  
961 HTFHRLEAGQYQIMIASVSGSLKNQINNVGRTVPASVQGVIAADNAYSYSLLIVSWQKAA 1020  
1021 GVAERYDILLTENGILLRNTSEPAITKOHKFEDLTPGKYYKIQLTVSGGLSKEAQTE 1080  
1021 GVAERYDILLTENGILLRNTSEPAITKOHKFEDLTPGKYYKIQLTVSGGLSKEAQTE 1080  
1081 GRTVPAAVTDLRTIENSTRHLSPFWTASGELSWMYNIFLYNPDGNLQERAQVDPVQSPS 1140  
1081 GRTVPAAVTDLRTIENSTRHLSPFWTASGELSWMYNIFLYNPDGNLQERAQVDPVQSPS 1140  
1141 FQNLQGRMYKMWIVTHSGELSNESFIFGRTPVASVSHLRGSRNRTTDSLFWNWPASGD 1200  
1141 FQNLQGRMYKMWIVTHSGELSNESFIFGRTPVASVSHLRGSRNRTTDSLFWNWPASGD 1200  
1201 PDFYELLINPNCTKKNWKKDLTEWRFGQLVGRKYVLWVTHSGDLNKNKVTASRTA 1260

1201 FDFYELLINPNCTKKNWKKDLTEWRFGQLVGRKYVLWVTHSGDLNKNKVTASRTA 1260  
1261 PSPSPSLMSFADIANTSLAITWKGPDPWTDYNDFELQWLPRDALTYFNPYNRSEGRIVY 1320  
1261 PSPSPSLMSFADIANTSLAITWKGPDPWTDYNDFELQWLPRDALTYFNPYNRSEGRIVY 1320  
1321 GLRPGRSYQNVKTVSGDSWKTYSKPIFGSVRTKPKDKIQNLHCRPQNSTAIACSWIPDPS 1380  
1321 GLRPGRSYQNVKTVSGDSWKTYSKPIFGSVRTKPKDKIQNLHCRPQNSTAIACSWIPDPS 1380  
1381 DFDGYSIECKRMDTQEVFESRLEKESLLNIMMLVPHKRYLSIKVQSGAGTSEVVEDS 1440  
1381 DFDGYSIECKRMDTQEVFESRLEKESLLNIMMLVPHKRYLSIKVQSGAGTSEVVEDS 1440  
1441 TITMIDRPPPPHPIRVNEKDVLSKSSINFTVNCWFSDTNGAVKFTVVVREADSGDE 1500  
1441 TITMIDRPPPPHPIRVNEKDVLSKSSINFTVNCWFSDTNGAVKFTVVVREADSGDE 1500  
1501 LKPEQOHPPLPSYLEYRHNASIRVYQTNYPASKCAENPNSKSFNIKLGAEMLGGKRD 1560  
1501 LKPEQOHPPLPSYLEYRHNASIRVYQTNYPASKCAENPNSKSFNIKLGAEMLGGKRD 1560  
1561 PTQOKFCDCGPKPHTAYRISIRAFITQDFDLDKEFTKPLYSDTFFSLPITTESPLFGAI 1620  
1561 PTQOKFCDCGPKPHTAYRISIRAFITQDFDLDKEFTKPLYSDTFFSLPITTESPLFGAI 1620  
1621 EGVSAGLFLIGMLVAVVALLICQKVSHGRERPSARLSIRDRPLSVHLNLGQKGRKTS 1680  
1621 EGVSAGLFLIGMLVAVVALLICQKVSHGRERPSARLSIRDRPLSVHLNLGQKGRKTS 1680  
1681 CPKINQFEGHFHMKLOADSNYLLSKYEELKDVGRNQSCDIALPENRGNKNYNNILPYD 1740  
1681 CPKINQFEGHFHMKLOADSNYLLSKYEELKDVGRNQSCDIALPENRGNKNYNNILPYD 1740  
1741 ATRVKLSNVDDDDPCSDYINASYIPGNFRREYITVQGPLPGTKDDFKWVWEQVNHVIM 1800  
1741 ATRVKLSNVDDDDPCSDYINASYIPGNFRREYITVQGPLPGTKDDFKWVWEQVNHVIM 1800  
1801 VTQCVKGRVKCHYWPADQDSLYYGLIILQMLSESVLPEWTIREFKICGEQJDAHRLI 1860  
1801 VTQCVKGRVKCHYWPADQDSLYYGLIILQMLSESVLPEWTIREFKICGEQJDAHRLI 1860  
1861 RHFHYTWPDHGPVETQSLIQFVRTVDRVYNINSPGAGPTVHCSAGVGRGTGTIALDRI 1920  
1861 RHFHYTWPDHGPVETQSLIQFVRTVDRVYNINSPGAGPTVHCSAGVGRGTGTIALDRI 1920  
1921 LQQLDSKDSVDIYGAVHDLHLHRVHMVQTECQVYVYLHQCVRDVLRLARKLSEQENLPFI 1980  
1921 LQQLDSKDSVDIYGAVHDLHLHRVHMVQTECQVYVYLHQCVRDVLRLARKLSEQENLPFI 1980  
1981 YENVNPEYHRDPVYSRH 1997  
1981 YENVNPEYHRDPVYSRH 1997

## RESULT 3

US-09-909-567B-54  
; Sequence 54, Application US/09909567B  
; Publication No. US2003002257A1  
; GENERAL INFORMATION:  
; APPLICANT: Macina, Roberto A.  
; APPLICANT: Nair, Manoj  
; APPLICANT: Chen, Seiyu  
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes  
; FILE REFERENCE: DEX-0214  
; CURRENT APPLICATION NUMBER: US/09/909,567B  
; CURRENT FILING DATE: 2001-07-20  
; PRIOR APPLICATION NUMBER: 60/219,834  
; PRIOR FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 54  
; LENGTH: 1997

; TYPE: PRT									
; ORGANISM: Homo sapien									
US-09-567B-54									
Query Match 99.8%; Score 10460; DB 3; Length 1997;									
Best local similarity 99.8%; Pred. No. 0;									
Matches 1993; Conservative 2; Mismatches 2; Indels 0; Gaps 0;									
Qy	1	MLSHGAGLAWITLSLQTLGLAEPERCNFTLAESKASSHSVSIQWRILGSPCNFSLIYSS	60						
Db	1	MLSHGAGLAWITLSLQTLGLAEPERCNFTLAESKASSHSVSIQWRILGSPCNFSLIYSS	60						
Qy	61	DTLGAALCPTFRIDNTTYGNCNLQDLQAGTIYFNKIIISLDEERTVVLQTDPLPPARFGVSK	120						
Db	61	DTLGAALCPTFRIDNTTYGNCNLQDLQAGTIYFNRIISLDEERTVVLQTDPLPPARFGVSK	120						
Qy	121	EKTTSTGLHWMTTPSSGKVTSYEVQLFDENNOKIQGVQIOESTSWNEYTFFNLTAGSKYN	180						
Db	121	EKTTSTSLHWMTTPSSGKVTSYEVQLFDENNOKIQGVQIOESTSWNEYTFFNLTAGSKYN	180						
Qy	181	IATAVSGGRKRSVSYTNGSTVPSPVKDIGISTKANSLIISWSHSGNVERYRLMLMDKG	240						
Db	181	IATAVSGGRKRSVSYTNGSTVPSPVKDIGISTKANSLIISWSHSGNVERYRLMLMDKG	240						
Qy	241	ILVHGGVVDKHTSYAFHGLSPGYLYNLTVMTEAAGLQNYRWKLVRTAPMEVSNLKVTD	300						
Db	241	ILVHGGVVDKHTSYAFHGLSPGYLYNLTVMTEAAGLQNYRWKLVRTAPMEVSNLKVTD	300						
Qy	301	GSLSLTKVKVQRPNGVDVSNITLSHGKTTIKESRVLAPWITETHFKELVPGRLYQVTVSC	360						
Db	301	GSLSLTKVKVQRPNGVDVSNITLSHGKTTIKESRVLAPWITETHFKELVPGRLYQVTVSC	360						
Qy	361	VSGELSAQKAVAGRTFPDKVANLEANNNGRMRSILVSVSWSPAGDWEQYRILLFNDSVLL	420						
Db	361	VSGELSAQKAVAGRTFPDKVANLEANNNGRMRSILVSVSWSPAGDWEQYRILLFNDSVLL	420						
Qy	421	NITVGEETOYVMDTGLVPGROYEVEVIVESGNLKNSERCQGRTPVLAQLVRKHANE	480						
Db	421	NITVGEETOYVMDTGLVPGROYEVEVIVESGNLKNSERCQGRTPVLAQLVRKHANE	480						
Qy	481	TSLSIMWQTPVAWEKIIISLARDLIIHKSLSKAKBFTFDLVPGRYKMATVTSISG	540						
Db	481	TSLSIMWQTPVAWEKIIISLARDLIIHKSLSKAKBFTFDLVPGRYKMATVTSISG	540						
Qy	541	DLKNSSSVKGRTPVPAQVTDLHVANQGMTSSLFTNWTQAOQGVDFYQVLLIHENNVILKNES	600						
Db	541	DLKNSSSVKGRTPVPAQVTDLHVANQGMTSSLFTNWTQAOQGVDFYQVLLIHENNVILKNES	600						
Qy	601	ISSETSRYSFSLKSGSLYVWVTVTSSGGISSRQVWVVEGRTPVSSVSGVTWNSGRNDYL	660						
Db	601	ISSETSRYSFSLKSGSLYVWVTVTSSGGISSRQVWVVEGRTPVSSVSGVTWNSGRNDYL	660						
Qy	661	SVSNLWAPGDVNDYEVTLSDHGKVVQSLVITAKSVRECSFSSLTPGRLYTVTITRSGKYE	720						
Db	661	SVSNLWAPGDVNDYEVTLSDHGKVVQSLVITAKSVRECSFSSLTPGRLYTVTITRSGKYE	720						
Qy	721	NHSFQERTVPDKVQGVSVNSARSQVLYRSVWHATGDFDHYETIKNNKNTQTSIPK	780						
Db	721	NHSFQERTVPDKVQGVSVNSARSQVLYRSVWHATGDFDHYETIKNNKNTQTSIPK	780						
Qy	781	SENECVFQVLPGRLYSVTVTTKSGQVEANEQNGRTIPFPVKDLTLRNRSTEDLHVTVS	840						
Db	781	SENECVFQVLPGRLYSVTVTTKSGQVEANEQNGRTIPFPVKDLTLRNRSTEDLHVTVS	840						
Qy	841	GANGDVQYIEIQLLFNDMKVFPFPHLVNTATEYRFTSLTPGRQYKILVLITISGVDQOSAF	900						
Db	841	GANGDVQYIEIQLLFNDMKVFPFPHLVNTATEYRFTSLTPGRQYKILVLITISGVDQOSAF	900						
Qy	901	IEGFTVPSAVKNTHISPNGATDSITVNWTPGGDVSQVTSVAPRHSQKVDQSQTIPKHVPE	960						
Db	901	IEGFTVPSAVKNTHISPNGATDSITVNWTPGGDVSQVTSVAPRHSQKVDQSQTIPKHVPE	960						
Qy	961	HTFHRLEAGRQYQIMIASVSGSLKNQINNVGRTVPASVQGVADNAYSSYSLIVSWQKAA	1020						

```
US-10-408-765A-2135
; Sequence 2135, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2135
; LENGTH: 1997
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2135

Query Match          99.8%; Score 10460; DB 4; Length 1997;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1992; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1  MLSHGAGLAWITLSLQTLGLAEPERCNFTLAESKASHSVSIOWRILGSPCNFSLIYSS 60
Db 1  MLSHGAGLAWITLSLQTLGLAEPERCNFTLAESKASHSVSIOWRILGSPCNFSLIYSS 60

Qy 61  DTLGAALCPTFRIDNTTYGNCLODQAGTIYFNFKIISLDEERTVVLQDPLPPARFVSK 120
Db 61  DTLGAALCPTFRIDNTTYGNCLODQAGTIYFNFKIISLDEERTVVLQDPLPPARFVSK 120

Qy 121  EKTISTGLHWWTSSGKVTSEYVOLFENNOKIQGOVIOESTSWNEYTFNLTAGSKYN 180
Db 121  EKTISTGLHWWTSSGKVTSEYVOLFENNOKIQGOVIOESTSWNEYTFNLTAGSKYN 180

Qy 181  IAITAVSGGKRSFVYNTGTVSPVKDIGISTKANSLLSWSHGSGNVERYRLMLMDKG 240
Db 181  IAITAVSGGKRSFVYNTGTVSPVKDIGISTKANSLLSWSHGSGNVERYRLMLMDKG 240

Qy 241  ILVHGGVVDKHTASYAFHGLSPGYLYNLVWTEAAGLQNYRWKLVRTPAPMEVSNLKYTND 300
Db 241  ILVHGGVVDKHTASYAFHGLTPGYLYNLVWTEAAGLQNYRWKLVRTPAPMEVSNLKYTND 300

Qy 301  GSLTSLKVKWQRPNGVDSYNITLSHGKTIKESRVLA PWITETHFKELVGRLYQVTVSC 360
Db 301  GSLTSLKVKWQRPNGVDSYNITLSHGKTIKESRVLA PWITETHFKELVGRLYQVTVSC 360

Qy 361  VSGELSAQKMAVGRTFPDKVANLEANNRMRSLVSVSPAGDWEQYRILLFNDSVLL 420
Db 361  VSGELSAQKMAVGRTFPDKVANLEANNRMRSLVSVSPAGDWEQYRILLFNDSVLL 420

Qy 421  NITVGKEETOVMDDTGLVPGROYEVEVIVESGNLKNSERCQGRTPVLA VLQLRVGHANE 480
Db 421  NITVGKEETOVMDDTGLVPGROYEVEVIVESGNLKNSERCQGRTPVLA VLQLRVGHANE 480

Qy 481  TSLSIMMQTPVAWEKIIISLADRDLLIIHKSLSKDAKEFTFDLVPGRYMATVTSISG 540
Db 481  TSLSIMMQTPVAWEKIIISLADRDLLIIHKSLSKDAKEFTFDLVPGRYMATVTSISG 540

Qy 541  DLKNSSSVKGRTPVPAQVTDLHVANQGMTSSIFTNWTOAQGDVFYQVLLIHENVVIKNES 600
Db 541  DLKNSSSVKGRTPVPAQVTDLHVANQGMTSSIFTNWTOAQGDVFYQVLLIHENVVIKNES 600

Qy 601  ISSETSRYSHLSKSGSLYSVVTTVSSGGISSRQVVVGGRTVPSSVSGVTVNNSGRNDYL 660
Db 601  ISSETSRYSHLSKSGSLYSVVTTVSSGGISSRQVVVGGRTVPSSVSGVTVNNSGRNDYL 660

Qy 661  SVSWLAPAGVDVNYEVTLSHDGKVQSLVIAKSVRECSFSLTPGRLYTVTTITRSKYE 720
Db 661  SVSWLAPAGVDVNYEVTLSHDGKVQSLVIAKSVRECSFSLTPGRLYTVTTITRSKYE 720
```

Db 1741 ATRVKLSNVDDPCSDYINASYIFPGNFRREYIVTQGLPGTKDDFWQVMEQVNHVIM 1800  
Qy 1801 VTQCVKGRVKCDHYWADODSLYYGDLILQMLSESVLPEWTIREFKICGEEQLDAHRLI 1860  
Db 1801 VTQCVKGRVKCDHYWADODSLYYGDLILQMLSESVLPEWTIREFKICGEEQLDAHRLI 1860  
Qy 1861 RHPHYTWPDHGVPETTTQSLIQFVTRVDRDYINRSPGAGPTVHCSAGVGRGTGFIALDRI 1920  
Db 1861 RHPHYTWPDHGVPETTTQSLIQFVTRVDRDYINRSPGAGPTVHCSAGVGRGTGFIALDRI 1920  
Qy 1921 LQQLDSKSDVDIYGAVHDLRLHRVHMVQTECOYVYLHQCVRDVLARKLSBOENPLFPI 1980  
Db 1921 LQQLDSKSDVDIYGAVHDLRLHRVHMVQTECOYVYLHQCVRDVLARKLSBOENPLFPI 1980  
Qy 1981 YENNVPEYHRDPVYSRH 1997  
Db 1981 YENNVPEYHRDPVYSRH 1997  
RESULT 5  
US-10-497-692-4  
; Sequence 4, Application US/10497692  
; Publication No. US2005004056A1  
; GENERAL INFORMATION:  
; APPLICANT: Meise, Martin  
; APPLICANT: Eulenberg, Karsten  
; APPLICANT: Fritsch, Rudiger  
; APPLICANT: Hader, Thomas  
; APPLICANT: Bronner, Gunter  
; APPLICANT: Steuernagel, Arnd  
; TITLE OF INVENTION: PTP10D, Tec protein tyrosine kinase and E2TP homologous proteins  
; TITLE OF INVENTION: involved in the regulation of energy homeostasis  
; FILE REFERENCE: 2923-632  
; CURRENT APPLICATION NUMBER: US/10/497,692  
; CURRENT FILING DATE: 2004-06-04  
; PRIOR APPLICATION NUMBER: PCT/EP02/13744  
; PRIOR FILING DATE: 2002-12-04  
; PRIOR APPLICATION NUMBER: EP 01 000 010.5  
; PRIOR FILING DATE: 2002-01-02  
; PRIOR APPLICATION NUMBER: EP 01 129 138.2  
; PRIOR FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: EP 01 128 844.6  
; PRIOR FILING DATE: 2001-12-04  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4  
; LENGTH: 1997  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-497-692-4  
Query Match 99.8%; Score 10460; DB 5; Length 1997;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1992; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 MLSHGAGLAWITLSLQTLGAEPCNFTLAESKASSHSVSIQWRILGSPCNFSLIYSS 60  
Db 1 MLSHGAGLAWITLSLQTLGAEPCNFTLAESKASSHSVSIQWRILGSPCNFSLIYSS 60  
Qy 61 DTLGAALCPTFRIDNTTYGNCNLDQAGTIYFNFKIISLDEERTVVLQTDPLPPARFVSK 120  
Db 61 DTLGAALCPTFRIDNTTYGNCNLDQAGTIYFNFKIISLDEERTVVLQTDPLPPARFVSK 120  
Qy 121 EKTSTGLHWVWTPSSGKVTSYEVQLFDENNQKIQGVQIOESTSWNEYTFNLTAGSKYN 180  
Db 121 EKTSTSLHWVWTPSSGKVTSYEVQLFDENNQKIQGVQIOESTSWNEYTFNLTAGSKYN 180  
Qy 181 IAITAVSGGRKRSFVYTNSTGTPSPVKDIGISTKANSLLSWHSNGNVERYRLMLMDKG 240  
Db 181 IAITAVSGGRKRSFVYTNSTGTPSPVKDIGISTKANSLLSWHSNGNVERYRLMLMDKG 240  
Qy 241 ILVHGGVVDKHTSYAFHGLSPGYLYNLTVMTEAAGLQNYRWKLVRTAPMEVSNLKVTDND 300

Db 241 ILVHGGVVDKHTSYAFHGLTPGYLYNLTVMTEAAGLQNYRWKLVRTAPMEVSNLKVTDND 300  
Qy 301 GSJTSLSUKVKQRPNGVDSVNIITLSHKGTIKESRVLAPWITETHEFKELVPGRLYQVTVSC 360  
Db 301 GSJTSLSUKVKQRPNGVDSVNIITLSHKGTIKESRVLAPWITETHEFKELVPGRLYQVTVSC 360  
Qy 361 VSGELSAQKMAVGRTTPDKVANLEANNNGRMRLSVVSWSPAGDWBEOYRILLFNDSVLL 420  
Db 361 VSGELSAQKMAVGRTTPDKVANLEANNNGRMRLSVVSWSPAGDWBEOYRILLFNDSVLL 420  
Qy 421 NITVGEETOYVMDTGLVPGRQYEVFEVIVESGNLKNSERCQGRTPVLAVLQVRKHANE 480  
Db 421 NITVGEETOYVMDTGLVPGRQYEVFEVIVESGNLKNSERCQGRTPVLAVLQVRKHANE 480  
Qy 481 TSLSIMWOTPVAEWKYYISLADRDLLTHKSLSKDAKEFTFTDLVPGRYMATVTSISG 540  
Db 481 TSLSIMWOTPVAEWKYYISLADRDLLTHKSLSKDAKEFTFTDLVPGRYMATVTSISG 540  
Qy 541 DLKNSSSVKGRTPVPAQVTDLHVANQGMTSSLFTNWTQAQGDVEFYQVLLIHENVVKNES 600  
Db 541 DLKNSSSVKGRTPVPAQVTDLHVANQGMTSSLFTNWTQAQGDVEFYQVLLIHENVVKNES 600  
Qy 601 ISSETSRYSFHSLSKSGSLYSVVVTVSGGISRSQVVGRTVPSSVSGVTNNSGRNDYL 660  
Db 601 ISSETSRYSFHSLSKSGSLYSVVVTVSGGISRSQVVGRTVPSSVSGVTNNSGRNDYL 660  
Qy 661 SVSWLVA PGDVNVEYTLSDHGKVQSLVIAKSVRECSFSLTPGRLYTVTITRSGKYE 720  
Db 661 SVSWLVA PGDVNVEYTLSDHGKVQSLVIAKSVRECSFSLTPGRLYTVTITRSGKYE 720  
Qy 721 NISFSOERTVPDKVQGVSVNSARSVDYLRVSWVHATGDFDHYEVTIKNNFNITQTSIPK 780  
Db 721 NISFSOERTVPDKVQGVSVNSARSVDYLRVSWVHATGDFDHYEVTIKNNFNITQTSIPK 780  
Qy 781 SENECEVQVQLVPGRLYSVTVTKSGQYEAHQNGRTIPEPVKDLTLNRSTEDLHVTVS 840  
Db 781 SENECEVQVQLVPGRLYSVTVTKSGQYEAHQNGRTIPEPVKDLTLNRSTEDLHVTVS 840  
Qy 841 GANGDVDOYEIOLLFNDKMFPPFHLVNTATEYRFTSLTPGRQYKILVLTISGDVQSAF 900  
Db 841 GANGDVDOYEIOLLFNDKMFPPFHLVNTATEYRFTSLTPGRQYKILVLTISGDVQSAF 900  
Qy 901 IEGETVPSAVKXNIHISPNGATDSLTVNWTFGGDDVDSYTVSAPRHSQKVDSDQTIKRVFE 960  
Db 901 IEGETVPSAVKXNIHISPNGATDSLTVNWTFGGDDVDSYTVSAPRHSQKVDSDQTIKRVFE 960  
Qy 961 HTFHRLEAGEQYQIMIASVSGSLKNQINNVGRTVPASVQGVIAADNAYSSYSLIVSQKAA 1020  
Db 961 HTFHRLEAGEQYQIMIASVSGSLKNQINNVGRTVPASVQGVIAADNAYSSYSLIVSQKAA 1020  
Qy 1021 GVAERYDIIILLTENGILLRNTSEPAATTKOHPEDLTPGKKYKIQILTIVSGGLFSKQAOTE 1080  
Db 1021 GVAERYDIIILLTENGILLRNTSEPAATTKOHPEDLTPGKKYKIQILTIVSGGLFSKQAOTE 1080  
Qy 1081 GRTVPAVTDLRITENSTRHLSPRWTASBEGELSWNYIFLYNPDGNLQERAQVDPVQSFS 1140  
Db 1081 GRTVPAVTDLRITENSTRHLSPRWTASBEGELSWNYIFLYNPDGNLQERAQVDPVQSFS 1140  
Qy 1141 FQNLQGRMYKMWIVTHSGELSNESFI FGRTVPASVSHLRGSRNRTTDSLFWNWSPASGD 1200  
Db 1141 FQNLQGRMYKMWIVTHSGELSNESFI FGRTVPASVSHLRGSRNRTTDSLFWNWSPASGD 1200  
Qy 1201 FQFYEYLILNPNGTCKENWKDLDTEWRFGGLVPGRYKYLWVVTTHSGDLSNKVTASRTA 1260  
Db 1201 FQFYEYLILNPNGTCKENWKDLDTEWRFGGLVPGRYKYLWVVTTHSGDLSNKVTASRTA 1260  
Qy 1261 PGPPLSLSFADIANSTSLAITWKGGPDDTDYNDPELQWLPRDALTVFNPYNNRSEGRIVY 1320  
Db 1261 PGPPLSLSFADIANSTSLAITWKGGPDDTDYNDPELQWLPRDALTVFNPYNNRSEGRIVY 1320  
Qy 1321 GLRPGRSYQFNKVTSGDSKWKYSKPIFGSVRTKPKIKNLHCRPQNSTAIACSWIPPPDS 1380  
Db 1321 GLRPGRSYQFNKVTSGDSKWKYSKPIFGSVRTKPKIKNLHCRPQNSTAIACSWIPPPDS 1380



1381 DFDGYSIECRQMDTOEVEFSSKLEKEKSLINIMLVPHKRYLVSIVKQOSAGTSEVVEDS 1440  
1381 DFDGYSIECRQMDTOEVEFSSKLEKEKSLINIMLVPHKRYLVSIVKQOSAGTSEVVEDS 1440  
1441 TITMIDRPPPPPHIRVNEKDVLSKSSINFVNCSWFSQDNGAVKFTVVVRADGSDE 1500  
1441 TITMIDRPPPPPHIRVNEKDVLSKSSINFVNCSWFSQDNGAVKFTVVVRADGSDE 1500  
1501 LKPEQOHPPLSYLYRNASIRVQTNVYFASKCAENPNNSKSNFIKLGAEMLSGKCD 1560  
1501 LKPEQOHPPLSYLYRNASIRVQTNVYFASKCAENPNNSKSNFIKLGAEMLSGKCD 1560  
1561 PTQOKFCGDKPKHTAYRISRAETQLEDEKFTKPLYSYDTFESLPITTESPLFGAI 1620  
1561 PTQOKFCGDKPKHTAYRISRAETQLEDEKFTKPLYSYDTFESLPITTESPLFGAI 1620  
1621 EGVSAAGFLIGMLVAVVALLICROKVGSHGRERPSARLSIRDRPLSVHLNLGQGNKRTS 1680  
1621 EGVSAAGFLIGMLVAVVALLICROKVGSHGRERPSARLSIRDRPLSVHLNLGQGNKRTS 1680  
1681 CPIKINQFEGHFMKLQADSNLYLSKEYEELKDVGRNQSCDIALLPENKGNRYNNILPYD 1740  
1681 CPIKINQFEGHFMKLQADSNLYLSKEYEELKDVGRNQSCDIALLPENKGNRYNNILPYD 1740  
1741 ATRVKLSNVDDPCSDINASYIIPGNFRREYIIVTQGLPGTKDDFKWVMEQNVHNI 1800  
1741 ATRVKLSNVDDPCSDINASYIIPGNFRREYIIVTQGLPGTKDDFKWVMEQNVHNI 1800  
1801 VTQCVKGRVKCDHYWADQSLYGLLQMLSESVLPEWTIREFKICGEEQDLDAHRLI 1860  
1801 VTQCVKGRVKCDHYWADQSLYGLLQMLSESVLPEWTIREFKICGEEQDLDAHRLI 1860  
1861 RHFHYTWPDHGVPETTQSLIQFVTRVDRYINRSPGAGTVVHCSAGVGTGTGFIADRI 1920  
1861 RHFHYTWPDHGVPETTQSLIQFVTRVDRYINRSPGAGTVVHCSAGVGTGTGFIADRI 1920  
1921 LQQLDSKSDVDIYGAVHDLRLHRVHMVQTECOYVYLHQCVRDVLARKLSEQNPLFPI 1980  
1921 LQQLDSKSDVDIYGAVHDLRLHRVHMVQTECOYVYLHQCVRDVLARKLSEQNPLFPI 1980  
1981 YENVNPEVHRDPVYSRH 1997  
1981 YENVNPEVHRDPVYSRH 1997

RESULT 6  
US-10-497-692-14  
; Sequence 14, Application US/10497692  
; Publication No. US20050004056A1  
; GENERAL INFORMATION:  
; APPLICANT: Meise, Martin  
; APPLICANT: Eulenberg, Karsten  
; APPLICANT: Fritsch, Rudiger  
; APPLICANT: Hader, Thomas  
; APPLICANT: Bronner, Gunter  
; APPLICANT: Steuernagel, Arnd  
; TITLE OF INVENTION: PTP10D, Tec protein tyrosine kinase and E2PT homologous proteins  
; TITLE OF INVENTION: involved in the regulation of energy homeostasis  
; FILE REFERENCE: 2923-632  
; CURRENT APPLICATION NUMBER: US/10/497,692  
; CURRENT FILING DATE: 2004-06-04  
; PRIOR APPLICATION NUMBER: PCT/EP02/13744  
; PRIOR FILING DATE: 2002-12-04  
; PRIOR APPLICATION NUMBER: EP 01 000 010.5  
; PRIOR FILING DATE: 2002-01-02  
; PRIOR APPLICATION NUMBER: EP 01 129 138.2  
; PRIOR FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: EP 01 128 844.6  
; PRIOR FILING DATE: 2001-12-04  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 14

LENGTH: 1450  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-497-692-14  
Query Match 72.6%; Score 7614; DB 5; Length 1450;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1448; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 519 EFTFTDLVPGKRYMATVTSISGDLKNSSSVKGRTVPAQVTDLHVANQMGTSLSLFTNNQ 578  
Db 1 EFTFTDLVPGKRYMATVTSISGDLKNSSSVKGRTVPAQVTDLHVANQMGTSLSLFTNNQ 60  
Qy 579 QGVDFEYQVLLIHENVVKNESISSETSRYSFSLKSGSLYSVVVTVTSGGISRQVVE 638  
Db 61 QGVDFEYQVLLIHENVVKNESISSETSRYSFSLKSGSLYSVVVTVTSGGISRQVVE 120  
Qy 639 GRTPVSSVSGVTNNNSGRNDYLSVSLVAPGDVDNVEVTLSDGKVVQSLVIAKSVRECS 698  
Db 121 GRTPVSSVSGVTNNNSGRNDYLSVSLVAPGDVDNVEVTLSDGKVVQSLVIAKSVRECS 180  
Qy 699 FSSLTPGRLYTVTITTRSGKYENHSFQERTVPDKVQGVSVNSARSADYLRVSWHATGD 758  
Db 181 FSSLTPGRLYTVTITTRSGKYENHSFQERTVPDKVQGVSVNSARSADYLRVSWHATGD 240  
Qy 759 FDHYEVTIKNNKNIQTKSIPKSENECVFQVLPGRLYSVTVTTKSGQYEAENEQNGRTI 818  
Db 241 FDHYEVTIKNNKNIQTKSIPKSENECVFQVLPGRLYSVTVTTKSGQYEAENEQNGRTI 300  
Qy 819 PEPVKDLTLNRSTEDLHVTWANGDGDVQYEQLLFNDMKVPPPHLVNTATERYTSL 878  
Db 301 PEPVKDLTLNRSTEDLHVTWANGDGDVQYEQLLFNDMKVPPPHLVNTATERYTSL 360  
Qy 879 TPCQRYKILVTISGVDQVQSAFIEGFTVPSAVKNIHISPNGATDSLTVNNTPGGVDVSY 938  
Db 361 TPCQRYKILVTISGVDQVQSAFIEGFTVPSAVKNIHISPNGATDSLTVNNTPGGVDVSY 420  
Qy 939 TVSAFHSQKVDSTQTPKHVFEHTFHRLEAGEQYQIMIASVSGSLKNOINNVGRTVPASV 998  
Db 421 TVSAFHSQKVDSTQTPKHVFEHTFHRLEAGEQYQIMIASVSGSLKNOINNVGRTVPASV 480  
Qy 999 QGVADNAYSYSILVSWQKAAQVABRYDILLTENGILLRNTSEPAATTKHGFEDLTPG 1058  
Db 481 QGVADNAYSYSILVSWQKAAQVABRYDILLTENGILLRNTSEPAATTKHGFEDLTPG 540  
Qy 1059 KKYKIQLITVSGGLFSKEAQTEGRTVPAAVTDLRITENSTRHLSFRWTASEGELSWNIF 1118  
Db 541 KKYKIQLITVSGGLFSKEAQTEGRTVPAAVTDLRITENSTRHLSFRWTASEGELSWNIF 600  
Qy 1119 LYNPDGNLQRAQVDPVLVQSFQNLQGRMYKXVIIVTHSGELSNESFIIGRTVPASVSH 1178  
Db 601 LYNPDGNLQRAQVDPVLVQSFQNLQGRMYKXVIIVTHSGELSNESFIIGRTVPASVSH 660  
Qy 1179 LRGSNRTTDSLWFNNSPASGDFDEYELILYNPNTGKKNKDKLTERFQGLVPGKRY 1238  
Db 661 LRGSNRTTDSLWFNNSPASGDFDEYELILYNPNTGKKNKDKLTERFQGLVPGKRY 720  
Qy 1239 VLWVTHSGDLSNKRVAESSTAPSPSLMSFADIANSTSLAITWKGPDPDWDYNDYDFELQWL 1298  
Db 721 VLWVTHSGDLSNKRVAESSTAPSPSLMSFADIANSTSLAITWKGPDPDWDYNDYDFELQWL 780  
Qy 1299 PRDALTVFNPYNNRKSEGRIVYGLRPGRSYQFNVTVSGDSWKTSYKPIFGSVRTKPKDKI 1358  
Db 781 PRDALTVFNPYNNRKSEGRIVYGLRPGRSYQFNVTVSGDSWKTSYKPIFGSVRTKPKDKI 840  
Qy 1359 QNLHCRPQNSTAIACSWIPDSDPDGYSIECRKMDTOEVEFSSKLEKEKSLINIMLVPH 1418  
Db 841 QNLHCRPQNSTAIACSWIPDSDPDGYSIECRKMDTOEVEFSSKLEKEKSLINIMLVPH 900  
Qy 1419 KRYLVSIKQOSAGMTSEVEDSTITMIDRPPPPPHIRVNEKDVLSKSSINFVNCSWF 1478  
Db 901 KRYLVSIKQOSAGMTSEVEDSTITMIDRPPPPPHIRVNEKDVLSKSSINFVNCSWF 960

Qy 1479 SDTNGAVKYFTVVVREADGSDDELKPEQOHPPLSPYLEYRHNASIRVYQNTYNFASKCAENPN 1538  
Db 961 SDTNGAVKYFTVVVREADGSDDELKPEQOHPPLSPYLEYRHNASIRVYQNTYNFASKCAENPN 1020  
Qy 1539 SNSKSNFKLGAEMESLGGKDDPTQKPCDGLKPHYATYRISIRAFQOLDEDLKEFTKP 1598  
Db 1021 SNSKSNFKLGAEMESLGGKDDPTQKPCDGLKPHYATYRISIRAFQOLDEDLKEFTKP 1080  
Qy 1599 LYSTDFSLPTTTESEPLFGAIEGVSAGLFLIGMLVAVALLICROKVSCHGRERPSARLS 1658  
Db 1081 LYSTDFSLPTTTESEPLFGAIEGVSAGLFLIGMLVAVALLICROKVSCHGRERPSARLS 1140  
Qy 1659 IRRDRPLSVHLNLGQGNKRTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQ 1718  
Db 1141 IRRDRPLSVHLNLGQGNKRTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQ 1200  
Qy 1719 CDIALPENRGNRYNNILPYDATRVKLSNVDDDDPCSDYINASYIPGNFRREYIVTQGP 1778  
Db 1201 CDIALPENRGNRYNNILPYDATRVKLSNVDDDDPCSDYINASYIPGNFRREYIVTQGP 1260  
Qy 1779 LPGTDDDFWKMVWEQNVHNIWMVTCQVEKGRVKCDHYWPAQDQSLYYGDLILQMLSESVL 1838  
Db 1261 LPGTDDDFWKMVWEQNVHNIWMVTCQVEKGRVKCDHYWPAQDQSLYYGDLILQMLSESVL 1320  
Qy 1839 PEWTIREFKICGEBQDLAHLRIHFHYTVMPDHGVPETTSQSLIQFVRTVVDYINRSPGAG 1898  
Db 1321 PEWTIREFKICGEBQDLAHLRIHFHYTVMPDHGVPETTSQSLIQFVRTVVDYINRSPGAG 1380  
Qy 1899 PTVVHCSAGVGRGTGFIALDRILQQLDSKSDVIYGAHVHDLRLHRVHMVQTEQYVYLHQ 1958  
Db 1381 PTVVHCSAGVGRGTGFIALDRILQQLDSKSDVIYGAHVHDLRLHRVHMVQTEQYVYLHQ 1440  
Qy 1959 CVRDVLRARK 1968  
Db 1441 CVRDVLRARK 1450

## RESULT 7

US-10-634-027-4  
; Sequence 4, Application US/10634027  
; Publication No. US20040077065A1  
; GENERAL INFORMATION:  
; APPLICANT: Procter & Gamble Company  
; APPLICANT: Evdokimov, Artem G  
; APPLICANT: Pokross, Matthew E  
; TITLE OF INVENTION: Three Dimensional Coordinates of HPTbeta  
; FILE REFERENCE: 9045M2  
; CURRENT APPLICATION NUMBER: US/10/634,027  
; PRIOR FILING DATE: 2003-08-04  
; PRIOR FILING DATE: 2002-09-25  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 4  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-634-027-4

Query Match 17.5%; Score 1830; DB 4; Length 336;  
Best Local Similarity 100.0%; Pred. No. 2e-108;  
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1662 DRPLSVHLNLGQGNKRTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDI 1721  
Db 1 DRPLSVHLNLGQGNKRTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDI 60  
Qy 1722 ALLPENRGNRYNNILPYDATRVKLSNVDDDDPCSDYINASYIPGNFRREYIVTQGPLPG 1781  
Db 61 ALLPENRGNRYNNILPYDATRVKLSNVDDDDPCSDYINASYIPGNFRREYIVTQGPLPG 120  
Qy 1782 TKDDFWKMVWEQNVHNIWMVTCQVEKGRVKCDHYWPAQDQSLYYGDLILQMLSESVLPEW 1841

Db 121 TKDDFWKMVWEQNVHNIWMVTCQVEKGRVKCDHYWPAQDQSLYYGDLILQMLSESVLPEW 180  
Qy 1842 TIREFKICGEBQDLAHLRIHFHYTVMPDHGVPETTSQSLIQFVRTVVDYINRSPGAGPTV 1901  
Db 181 TIREFKICGEBQDLAHLRIHFHYTVMPDHGVPETTSQSLIQFVRTVVDYINRSPGAGPTV 240  
Qy 1902 VHCAGVGRGTGFIALDRILQQLDSKSDVIYGAHVHDLRLHRVHMVQTEQYVYLHOCVR 1961  
Db 241 VHCAGVGRGTGFIALDRILQQLDSKSDVIYGAHVHDLRLHRVHMVQTEQYVYLHOCVR 300  
Qy 1962 DVLRARKLRSEQENPLFPFIYENVNPEYHRDPVYSRH 1997  
Db 301 DVLRARKLRSEQENPLFPFIYENVNPEYHRDPVYSRH 336

## RESULT 8

US-10-634-027-7  
; Sequence 7, Application US/10634027  
; Publication No. US20040077065A1  
; GENERAL INFORMATION:  
; APPLICANT: Procter & Gamble Company  
; APPLICANT: Evdokimov, Artem G  
; APPLICANT: Pokross, Matthew E  
; TITLE OF INVENTION: Three Dimensional Coordinates of HPTbeta  
; FILE REFERENCE: 9045M2  
; CURRENT APPLICATION NUMBER: US/10/634,027  
; PRIOR FILING DATE: 2003-08-04  
; PRIOR FILING DATE: 2002-09-25  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 7  
; LENGTH: 319  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-634-027-7

Query Match 16.1%; Score 1691; DB 4; Length 319;  
Best Local Similarity 99.4%; Pred. No. 1.4e-99;  
Matches 312; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1662 DRPLSVHLNLGQGNKRTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDI 1721  
Db 2 DRPLSVHLNLGQGNKRTSCPIKINQFEGHFMKLQADSNYLLSKEYEELKDVGRNQSCDI 61  
Qy 1722 ALLPENRGNRYNNILPYDATRVKLSNVDDDDPCSDYINASYIPGNFRREYIVTQGPLPG 1781  
Db 62 ALLPENRGNRYNNILPYDATRVKLSNVDDDDPCSDYINASYIPGNFRREYIVTQGPLPG 121  
Qy 1782 TKDDFWKMVWEQNVHNIWMVTCQVEKGRVKCDHYWPAQDQSLYYGDLILQMLSESVLPEW 1841  
Db 122 TKDDFWKMVWEQNVHNIWMVTCQVEKGRVKCDHYWPAQDQSLYYGDLILQMLSESVLPEW 181  
Qy 1842 TIREFKICGEBQDLAHLRIHFHYTVMPDHGVPETTSQSLIQFVRTVVDYINRSPGAGPTV 1901  
Db 182 TIREFKICGEBQDLAHLRIHFHYTVMPDHGVPETTSQSLIQFVRTVVDYINRSPGAGPTV 241  
Qy 1902 VHCAGVGRGTGFIALDRILQQLDSKSDVIYGAHVHDLRLHRVHMVQTEQYVYLHOCVR 1961  
Db 242 VHCAGVGRGTGFIALDRILQQLDSKSDVIYGAHVHDLRLHRVHMVQTEQYVYLHOCVR 301  
Qy 1962 DVLRARKLRSEQEN 1975  
Db 302 DVLRARKLRSEQEN 315

## RESULT 9

US-10-634-027-6  
; Sequence 6, Application US/10634027  
; Publication No. US20040077065A1  
; GENERAL INFORMATION:  
; APPLICANT: Procter & Gamble Company  
; APPLICANT: Evdokimov, Artem G

```

; APPLICANT: Pokross, Matthew E
; TITLE OF INVENTION: Three Dimensional Coordinates of HTPbeta
; FILE REFERENCE: 9045M2
; CURRENT APPLICATION NUMBER: US/10/634,027
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US 60/413,547
; PRIOR FILING DATE: 2002-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-634-027-6

Query Match      16.1%; Score 1690; DB 4; Length 312;
Best Local Similarity 100.0%; Pred. No. 1.6e-99;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1662 DRPLSVHLNLGQGNKRTSCPIKINQEGHEPMKLOADSNYLLSKEYBELKDVGRNQCSDI 1721
Db 1 DRPLSVHLNLGQGNKRTSCPIKINQEGHEPMKLOADSNYLLSKEYBELKDVGRNQCSDI 60

Qy 1722 ALLPENKGRYNNILPYDATRVKLSNVDDPCSDYINASYPGNFRREYIVTQGPLG 1781
Db 61 ALLPENKGRYNNILPYDATRVKLSNVDDPCSDYINASYPGNFRREYIVTQGPLG 120

Qy 1782 TKDQFWKQWQVHNVMTQCVKGRVKCDHYWADQSDLYGDLILQMLSESILPEW 1841
Db 121 TKDQFWKQWQVHNVMTQCVKGRVKCDHYWADQSDLYGDLILQMLSESILPEW 180

Qy 1842 TIREFKICGBQLDAHLIRHFHYTVPDGHVPTTQSLQFVTRVDYINRSPGAGTV 1901
Db 181 TIREFKICGBQLDAHLIRHFHYTVPDGHVPTTQSLQFVTRVDYINRSPGAGTV 240

Qy 1902 VHCAGVGRGTCTFALDRILQQLSDKSDVDIYGAVHDLRLHRVHMVOTECOYVYLHOCVR 1961
Db 241 VHCAGVGRGTCTFALDRILQQLSDKSDVDIYGAVHDLRLHRVHMVOTECOYVYLHOCVR 300

Qy 1962 DVLRAKLRSEQ 1973
Db 301 DVLRAKLRSEQ 312

RESULT 10
US-11-097-143-2076
; Sequence 2076, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23

; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2076
; LENGTH: 1647
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-2076

Query Match      14.8%; Score 1556.5; DB 6; Length 1647;
Best Local Similarity 30.3%; Pred. No. 6.7e-90;
Matches 474; Conservative 254; Mismatches 627; Indels 209; Gaps 56;

Qy 519 EFTFTDLVGRKY-----MATVTSISGLKNSSSVKGRTVPAQVTDLH 561
Db 102 EIOFSRALPGTKNFWLYTNFTHDMLTWVIT-----TAPPPNLS 146

Qy 562 VANQGMTSSTFTNWTQAGDVEFY--QVLLIHENVVKNES--ISSTSRYSFSLKSGS 617
Db 147 VQVRSGKNAIILSPPTQGSYTAFAKIVGLSEASSYNRTFQVNDNTFQHSVKELTPGA 206

Qy 618 LYSVVVTVSGGSSRRQVVVEGRTVPSVSGVTVNNSGRNDYLSVWL--VAPGDVNYE 675
Db 207 TYQVQAYTIYDGKESVAYTSRNFTTKPTKGTIWFPRNETTLLVLMQPPYAGIYTHYK 266

Qy 676 VTLSDHGDVQVSLVIAK-----SVRECSFSSLTGRLTYTITTRS--GKYENHSFSOERT 729
Db 267 VSIEPPDANDSVLYVEKEGPPGPAQAFKGLVPGRAYNISVQVMSDEISLPTTAQTRT 326

Qy 730 VPDKVQCVSVNS--ARSDYLRLSVWVHATG--DFDHYEVTIKNNFNFTQKSIKSENECV 786
Db 327 VPLRLNVTDRDFITNSFRVLWEAPKGISSEFDKYOVSVATPR--RQSTVPSRNEPVA 383

Qy 787 FVQL----VPGRLYSVTVTTKSGQYEA--NEQGNRTIPEPVKDL--TLNRSTEDLHVTWS 840
Db 384 FDFRDLAEPGKTFNVIKTVSGKVTSPATGDTLRLPLPVRNLRSINDDKNTMIITWE 443

Qy 841 GANDVDVQYIQLLFNDMKVPPFPFHLVNTATEFRFT--SLTPGQYKILVLTISGDVQQ-- 897
Db 444 -ADPASTQDEYRIVYHELETFNGDSTLTATDRFTLESLLPGRNYSLSVQAVSKOMESN 502

Qy 898 --SAFTEGFTVPSA-----VKNHISPGATGDSLTNNWTPGGGDVDS-----YTV----- 940
Db 503 ETSIFV--VTRPSPIEDLSIRM-----GLNISWK--SDVNSKQEQYEVLYSRNG 550

Qy 941 -SAPRHSQVDSQTI PKHVEFTFHRLEAGEQYQIMIASVSGSLKNNQVNVGRTVPAASVQ 999
Db 551 TSDLRTOKTKESSLVKN-----LQCGAGYELKVFVSHDLRSEPHAYFQAVYENPP 602

Qy 1000 GVIADNAYSSYSLIVSWQKAAGVAERYDIIULLTENGILLRETS-----PATTQKHF 1052
Db 603 RNMTIETVRSNSVLVHMSPPESGE-----FTEYSIRYRTDSEQQWVRLPSVRSTADI 655

Qy 1053 EDLTPGKYKIQLITVSGGLFSKEAQTEGRTVPAAVTDLRI TENSTRHLFRWTA SEGEL 1112
Db 656 TDMTKGEYTIQNTVSFGVESVPQEVWTVTPNPVSNIIQLVDSNITLFPKPGGRV 715

Qy 1113 SWYNIFLYNPDG-----NLOERAQVDPL--VQSFSEFQNLQGRMKXVITVHS--GELSN 1163
Db 716 ESYLKWPSDNPGRVQTKVNSKNSADSLSTVRVLIGELMPPGVQYKFDIQTTSYGILSG 775

Qy 1164 ESFIFORTVP---ASVSHLRGSRNRTDLSLFWNSP---ASGDFDFTYELLINPDNGYKE 1217
Db 776 ITSILYPRTMPLIQSDVVVWANGEKEDERTITLSTYPTPQSSSKFPIYRFLSGDAEIRDK 835

Qy 1218 ---NWKDKDLTEWRFOGLVPGKVVLMVWVTHSGDLSN--KYTAESRTAPSPPSLMSPADIA 1273
Db 836 KLANDTDRKVT---FTGLVPEGRLYNITVTVSGVASLPQRQDRLYPEEITQLHATNIT 892

Qy 1274 NTSILAITWKGPDPDWTNDPELOWLPRDALTVENPNYNNRSEKRI--VYGLRPGRSYQFNV 1332
Db 893 DTEISLRWDLPGK--EYNEDFIAYLTADNLLA-----QNMTRNEITISDLRPHRYTFTV 946

Qy 1333 KTVSG--DSWKYSKPIFGSVRTK---PDKIQMLHCRPQNSTAIACSWIPDSDFDG--- 1384
```

Db 947 VRSSTESSVLRSSPLSASFTTNEAVPGVERPHPTDQVPSEINFEWSLPSSEANGVIR 1006  
Qy 1385 -YSIECRKMDTOVEFSEKLEKESLNNMVLPHKBYLVSIVKQSA-GMTSEVVEDSTI 1442  
Db 1007 QFSIAYTNINLTDAGQDSESEAFGVKNLKGETYVFKIQAKTAIGFGPREYRQTM 1066  
Qy 1443 TMDRPPPPPHIRVNEKDVLSKSSINFVNCISWFSDTNGAVKFTVVVREADGSDLK 1502  
Db 1067 PIL---APPRATQVTEVYRSTTIQIRFKNYFSDQNGQVRYIIIVAE-----DDAK 1119  
Qy 1503 PEQOHPLPSLEYVHNASIRYQF--NYFASKCAENPNSKSNFIKLGAMESLGGKRD 1560  
Db 1120 NASGLEMPSLDQVSYSWMLPYQAIDPY-----PFENRSVEDFTIGTE-----NCD 1166  
Qy 1561 PTQOKFCGDLKPHATYRISIRATQOLFDEDLKEFTKPLSDTFFSLPITTESEPLEGAI 1620  
Db 1167 NHKIGYCNGPLKSGTYYRVKRAFT-----GADKFT-----DTAYSPPITQDQNT-SLI 1215  
Qy 1621 EGVSAGLFLICMLVAVVALLICROKVGSHRE-RPSARLSIRRRDRPLSVHLNLQKGNRKT 1679  
Db 1216 VAITVPLTII--LVLLVTLIFYKRRNNCKTKTKDSRANNNSLPDSV-----IEQNR-- 1266  
Qy 1680 SCPIKINOFEGHFMKLOADSNLYLSKEYBELKDVGRNQSCDIALLPENROKNNRNILPY 1739  
Db 1267 --PILIKNFAHYRLMSADSDRFSEBELKHVGRDQPCCTFADLPCNRPKNFTNIPY 1324  
Qy 1740 DATRVKLSNVDDPCSDIYNASIVPGNNPREVITVQGPLCTGKDFWKVWQVQNVHIV 1799  
Db 1325 DHSRFLQPVDDDEGSDIYNANVPGHNSPREFIVTQGPLHSTRDDPFWRMCWESNRAIV 1384  
Qy 1800 MVTQCEKGRVKCDHYWPADQDLSLYGDLIOLMSVLPWETIRBPKIC-GEQOLDADR 1858  
Db 1385 MLTRCFEKGREKCDQVWPNVTVPVYGDIKVQILNDSHYADWWTETMLCRGSEQ-----R 1440  
Qy 1859 LIRHFHYVWPDHGVPETTSLOQVTRVTRDYINRSPGFTVWHCSAGVGRGTGTFIALD 1918  
Db 1441 ILRHFFHTTWFDEGVNPPQTLVRFVRAFRDRIGAE--QRPIVVHCSAGVGRSGTFTILD 1498  
Qy 1919 RILQOLDSKSDIYGAHVHDLRLHRVHMVQTEQVYVYLHOCVDRDVLARKLRSEQENPLF 1978  
Db 1499 RILQINTSDYDIFGIVYAMRKERVMMVQTEQYICIHQCLLAVLEGK-----ENIVG 1552  
Qy 1979 PIVE 1982  
Db 1553 PARE 1556

RESULT 11  
US-10-087-684-40  
; Sequence 40, Application US/10087684  
; Publication No. US20040029116A1  
; GENERAL INFORMATION:  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: MacDougall, John R.  
; APPLICANT: Miller, Isabelle  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Stone, David J.  
; APPLICANT: Grosse, William M.  
; APPLICANT: Lepley, Denise M.  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Burgess, Catherine E.  
; APPLICANT: Casman, Stacie, J.  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Boldog, Ferenc L.  
; APPLICANT: Li, Li  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Tchernev, Velizar T.  
; APPLICANT: Vernet, Corine A.M.  
; APPLICANT: Zerhusen, Bryan D.

; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Gangolli, Esha A.  
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-214 CIP  
; CURRENT FILING DATE: 2003-03-10  
; PRIOR FILING DATE: 2003-03-10  
; PRIOR FILING DATE: 2000-11-29  
; PRIOR FILING DATE: 2000-11-29  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR FILING DATE: 2001-08-20  
; PRIOR FILING DATE: 2001-08-20  
; PRIOR FILING DATE: 2001-10-05  
; PRIOR FILING DATE: 2001-10-05  
; NUMBER OF SEQ ID NOS: 220  
; SOFTWARE: Curasequid version 0.1  
; SEQ ID NO 40  
; LENGTH: 1767  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-10-087-684-40  
  
Query Match 14.2%; Score 1485.5; DB 4; Length 1767;  
Best Local Similarity 29.2%; Pred. No. 2.6e-85;  
Matches 475; Conservative 269; Mismatches 687; Indels 197; Gaps 62;  
  
Qy 456 KISERCOGRTVPLAVLQLRV---KHANETSLSIMQWTPVAWEK---YIISLA-----D 503  
Db 44 QNSQOQKHFWMLVVGILTFIAQHANAADLVI--NVPNASSNANAFRIDYSPPGFPE 101  
Qy 504 RDLLHLSKSDAKEFTFTDLVPGRKY--MATVTSISGDLKSSSVKGRTPAQVTDLH 561  
Db 102 PNTIIPASDIGDKI---FSRALPGTEYNFWLYYTNSTHREQLTWITVNTTADPPANLS 158  
Qy 562 VANQGMTSLSLTNWTQAQGVPEFYVLLIHENVVKNESIS---SETSRYSFHSLSKGSGL 618  
Db 159 VQLRSKSAFITWRPPGSGRYSGFIRVLGLTDLPPERSYSLEGNETLQLSAKELTPGGS 218  
Qy 619 YSVVTVTVSGGISRRQVVVEGRTVPSSVSGVTNNSSGRNDYLSVWL--VAPGDVNDYEV 676  
Db 219 YQVQAYSVYQKESVATSRNFTTKPTCKFIWFRNRTLLVLVWQPPPPAGIYTHYRV 278  
Qy 677 TLSDHGVVQSLVIAK-----SVRECSFSLTPGRLYTVTITTRSGKYENHSFS-----Q 726  
Db 279 SITPDDAIQSVLYVEREGEPGPAQAQAFKGLVPGREYNISVQTVS---EDETS SVTTAR 335  
Qy 727 ERTVPDKVQGVSVNS-ARSDYLRLVSW--VHATGDPHDYEVITIKNNFIQTKSIKSEN 783  
Db 336 YLTVPERVLNVTDFEAYTTSSSPRVVRWEPRTYSEFDAYQVMLSTERRIF---NVPRAN 392  
Qy 784 -ECVFVQ-----LVPGRLYSVTYTKSGOYANEQNGRII-PEPVKDL--TLNRSTEDL 835  
Db 393 GDSVYFDYDPILEPGRTYEVVVKTIADNVNWPASGEVTLRPRPRSLGGLFDDRANA-L 451  
Qy 836 HVTWSCA-NGDVDOYBIQL--LFNDMKVFPFPF-----HLVNTATEYRFTSLTPGRQYKI 886  
Db 452 HISWEPATGRQDSYRISVHEQTNASEVPAPPPVAESQITTNLTETLDSLLAGRYLI 511  
Qy 887 LVLTISGDVQOQSAF-IEGFTVPESA--VKNIHISPNGATDSLTVNWTTPGGDV-----DSYT 939  
Db 512 AVQALSKGVASNASDITRYTRPAAPLIQELRSIDQO---LMSWR---SDVNSRODRYE 564  
Qy 940 VSAFRHSQKVDQOTIPKHVFEHTFHELEAGEQYQIMIASVSGSLKQINNVGRTV-PASV 998  
Db 565 VHYQRNGTR-EERTMATNETSLTIHYLHPGSGEVKVAISHGVRSSEPHSYFOAVFPKPP 623  
Qy 999 QGVADNADYSYSLIVSQKAAGVABERYDILLITENGILLRNTSEP-----ATTQKHK 1052

Db 624 QNLTLQVHTNL-VVLHQAPEGSD-----FSEYVRYETDASPWQRISGLHENEARI 675  
Qy 1053 EDLTPGKKYKIQILTVSGGLFSKQATGRTVPAVTDLRITENSTRHLSFRWTASEGEL 1112  
Db 676 KDMHYGERLYQVNTVSGVSPHLELNLNMTMPQPVNSVPLVDSRNLTLLEWPRPDGHV 735  
Qy 1113 SWYAFILNPNQNLQERAQVDPVQ-----SFSQNLQGRMYKM-VIVTHSGELS 1162  
Db 736 DFYTLKMWPTDE--EDRVFNKVTQLEDLSFSVRPIEDLSPGRVFVQOASNGIRS 793  
Qy 1163 NESIFGRTVP-----ASVHLRGSNRNRTDLSLFWNWPASGD---DFEYELIILNYP 1211  
Db 794 GTTHLSFTYMLQSDVFIANAGHEQODEIT-----LSYTPPADSTFRFDIYFSMGDP 849  
Qy 1212 NGTKENWKDILTEWRFOGLVPGKRYVLMVWVTHSGDLSN-KVTAESRTAPSPSLMSFA 1270  
Db 850 TIKDKEKLANDTERKLSFGLTPGKLYNVVTVWTVSGVVASLPVQRLVRLHPLPISDLKAI 909  
Qy 1271 DIANTSIAITWKGPPDWDYNDPELOLPRDALTVFNPYNNRKSEGRIVYGLRPGRSYOF 1330  
Db 910 QVAAREITLHTWAPAG--EYTDLFQYLSADEEAPQLLQNTVKNTEITLQGLRPHYNYTF 967  
Qy 1331 NVKTVSGDSWKT-----YSKPIFGSVRT---KPKIQNLHCRPQNSTAIACSW 1375  
Db 968 TVVVRSGSIQTDVADSVSVTLMRSSAPISASYQTLTAPPGKVDYFQPSDVQPGEVTFEW 1027  
Qy 1376 -IPP---DSDFDGYSIIECRKMDTQEVFSRKLKEKSLNLMVLPHKRYLVSIKVQSA- 1430  
Db 1028 SLEPAEQHPIDYFRITQONADDAADVSSYFPPVNAVQTKIDGLVPGNHVIFRQAKSAL 1087  
Qy 1431 GMTSEVEDSTIWMIDRPPPPPHIRVNEKDVLSKSSINTVNCSPFSDNGAVKVFV 1490  
Db 1088 GYGAEREHIQMPIL---APVPPEFSTPLEVSTSTSIETISFRQGYFSAHGVRSYTI 1144  
Qy 1491 VVREAGSDLEKPEQHPPLPSVLEYRHNASTRVQTNVFAKCAE--NP-----NSNKSFP 1544  
Db 1145 IIAEDVGKNASGLE---MPSQD-----VQAY-TWLPVQALPEYNPFLTSGSRKS- 1192  
Qy 1545 NIKLGAEMESLG-GKRDPQOQKFCGDKPLKPHATYRISIRAFQTFDDELKFTKPLXSDT 1603  
Db 1193 --SLEAEHFTGTANCCKHQAGYCNGLRAGTTRYIKIRAPT---DED-----KFTDT 1240  
Qy 1604 FFSLPITTESPLEFAGIEGVSAGLFLIGMLVAVVALLICROKVSGRHRPRLSIRDR 1663  
Db 1241 VYSSPITTERSDTVIATVSA-VLLVAMVLVW---YCQHRQ--LIRRAKSLARQDE 1294  
Qy 1664 PLSVHLNLGQGNKRTSCPIKINOPEGHFMKLDQSDSNVLLSKYEELKQVGRNOSCDIAL 1723  
Db 1295 LAAL-----PEGYITPNRPHVKDFSEHYRIMSADSDFRFSEEFELKHGVRDQACSFAN 1349  
Qy 1724 LPENRGKNRYNNILPYDATTRVKSNDVDDPCSDYINASYIPGNFRPREYIVTQGLPGLTK 1783  
Db 1350 LPCNRPNRFTNLLPYDHSRFLQVDDDDSGSDYINANYMFGHNSPREFIVTQGLHSTR 1409  
Qy 1784 DDFWQWQVQNNHNVMTVQCVEKGRVKCDHYWADQSDLSYGLDILLOMSESVPBWTI 1843  
Db 1410 BEFRMCWESNSRAIVMLTRCFEGRKCDQYVVDVAMFYDGIKVQLIIDTHYDWSI 1469  
Qy 1844 REPKI---CGEQLDAHRLIBHFHYWPDHGVETTSQSLQFVRTVTDYINRSPGAGPT 1900  
Db 1470 SEFWVRNRC-----ESRIMHFHFTWPDGVPPEPQSLVRFVRAFRDVIQTD--MRPI 1521  
Qy 1901 VVHCSAGVGRGTGTIALDRILQQLDSDSDVDIYGAVHDLRLHRYVMVQTECQYVYLHCQV 1960  
Db 1522 IVHCSAGVGRSGTIALDRILQHLHKSVDYDIFGIVFAMRKERVFMVQTECQYVCIHQCL 1581  
Qy 1961 RDVLRAR 1968  
Db 1582 LAVLEGKE 1589

US-10-218-779-40  
; Sequence 40, Application US/10218779  
; Publication No. US20040029222A1  
; GENERAL INFORMATION:  
; APPLICANT: Edinger, Shlomit  
; APPLICANT: MacDougall, John  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Stone, David  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Alsbrook II, John  
; APPLICANT: Leplev, Denise  
; APPLICANT: Rieger, Daniel  
; APPLICANT: Burgess, Catherine  
; APPLICANT: Casman, Stacie  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Li, Li  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shenoy, Suresh  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Vernet, Corine  
; APPLICANT: Zernusen, Bryan  
; APPLICANT: Malyankar, Uriel  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Miller, Charles  
; APPLICANT: Gangolli, Beha  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-214  
; CURRENT APPLICATION NUMBER: US/10/218,779  
; CURRENT FILING DATE: 2002-08-14  
; PRIOR APPLICATION NUMBER: 60/253,834  
; PRIOR FILING DATE: 2000-11-29  
; PRIOR APPLICATION NUMBER: 60/250,-926  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: 60/264,180  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 60/313,656  
; PRIOR FILING DATE: 2001-08-20  
; PRIOR APPLICATION NUMBER: 60/327,456  
; PRIOR FILING DATE: 2001-10-05  
; NUMBER OF SEQ ID NOS: 216  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 40  
; LENGTH: 1767  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-10-218-779-40

Query Match 14.2%; Score 1485.5; DB 4; Length 1767;  
Best Local Similarity 29.2%; Pred. No. 2.6e-85;  
Matches 475; Conservative 269; Mismatches 687; Indels 197; Gaps 62;  
Qy 456 KNSERCQGRVPLAVLQLRV---KHANETSLSIMMOTFVAEWEK---YIISLA-----D 503  
Db 44 QNSQQQKHFWLWVGILTFIAQHANAADLVI--NVPNASSNANAFYRIDYSPFPGPPE 101  
Qy 504 RDLIIHKSUSKAKEFTTDLVPGRKY--MATVTSISGDLKNSSSVKGRTPVPAQVTDLH 561  
Db 102 PNTTIPASDICKOIK---FSRALPGTEYNFWLYTNSHREQLTWTNITTAPPPANLS 158  
Qy 562 VANQGMTSSLFTNTWQAGDVEFYQVLLIHENVIKNESIS---SETSRYSFHLKSGSL 618  
Db 159 VQLRSSAFITWRPPGSGRYSGRIRVLGTLDFPERSYSLEGNETLQLSAKELTPGGS 218  
Qy 619 YSVVVTTVSGISRRQVWVEGRTPVSSVGVTVNNSGRNDYLSVSL--VAPGVDVNYEV 676  
Db 219 YQVQAYSVOQKESVATSRNFTKPTKPGKFIWPRNETLLVLWQPPPPAGIYTHRV 278



QY 562 VANQWTSSTLFTNWTQAGDVFEYOVLLIHENVLIKESIS---SETSRYSFHSLSKGSGL 618  
Db 159 VOLRSSKAFITWRPPGSGRISGFRIRVLGLTDLFPERSYSLEGNETIQLSAKETLPGGS 218  
QY 619 YSVVVVTVSGGSSQVVVEGRTPSSVSGVTNNSGRNDYLSVSL--VAPGDVDNYEV 676  
Db 219 YQVQAYSVVQKESVAYTSRNPFTKPNTPGKFIWFRNETTLLVLWQPPFAGIVTHRV 278  
QY 677 TSHDGKVVQSLVIAK-----SVRECSFSSLTTPGRLYTVTTTSGKGYENHSFS-----Q 726  
Db 279 SITPDAAIQSVLYVEREGEPGPAQAAPKGLVPGREYNISQVTVS---EDETSVPPTAR 335  
QY 727 ERTVPDKVQGVSVNS-ARSDYLRVSW--VHATGDFDHVEYTIKKNFIQTKSPKEN 783  
Db 336 YLTVPERVLNVTDFDAYTSSFRVRWEPRTYSEFDAYQWLSSTRIP---NVPRAN 392  
QY 784 -ECVFPVQ---LVPGRLSVTVTTKSGQYEAENGRTI-PEPVKOL---TLNRKSTBDL 835  
Db 393 GDSVYFDYDILPGRTYEVVVVKTADNVNSWPASGEVTLRPPVRSLGGLDORSNA-L 451  
QY 836 HVTWSGA-NGVDVQVEIOL--LFDNMKVPPPP-----HLVNTATEYFTSLTPGRQYKI 886  
Db 452 HISMEPAETGRQDSYRISHEQTNASEVPAPPPVAABESQITNLTEYTLDSLLAGRYLI 511  
QY 887 LVLITSGDVQOSAF-IEGFTVPSA--VKNHISPNGATDSLTVNWTPOGGDV---DSYF 939  
Db 512 AVQALSKVASNADITRYTPAAPLIQELRSIDQG---LMLSWR---SDVNSRQDRYE 564  
QY 940 VSAFRHSOKVDSQTIKPIHVEFTHRLRAGEQYQIMIASVSGSLKNQINNVGRTV-PASV 998  
Db 565 VHYQENGTR-EERTWATNETSI.THYLHPGSGYEVKVHAI SHGVRSSEPHSFQVAPFKPP 623  
QY 999 QGVADNAYSSVSLIVSQKAGVAERYDIILLTENGILLRNTSEP-----ATTQKHK 1052  
Db 624 QNLTLQTVHTNL-VVLHWAQEGSD-----FSEYVVRXRTDASPQWRISGLHENEARI 675  
QY 1053 EDLTPGKKYKQILTVSGSLSKAEQTEGRTPAAVTDLRITENSTRHLSFRWTASEGEL 1112  
Db 676 KDMHYGERLYQVQNTVSGVSEPHLELNTMPQPVSNVNVPLVDSRLNLTLEWPRPDGHV 735  
QY 1113 SWNIFLYNPQNLQERQAVDPLVQ-----SFSQNLQGRMYKM-VIVTHSGLS 1162  
Db 736 DFYTLKWPTDE--EDRVEFNVTQLEDLSFSVRIPIEDLSFGQYFQVQASNGIRS 793  
QY 1163 NESFIFGRTP-----ASVSHLRGSRNNTDLSLWFWSPASGD---PDFYELILYNP 1211  
Db 794 GTTHLSTRTPMLIQSDVFIANAGHEQGDETIT---LSYTPPADSTRFDIYRFSMGDP 849  
QY 1212 NGTKENWKDLTEWRQGLVPRGKYVLVWVTHSGDLN-KVTAESRTAFSPSLMSFA 1270  
Db 850 TIKOKEKLANDTERKLSFSGLTPGKLYNVTVMTVSGVASLPVQRLYRHLPLISDLKAI 909  
QY 1271 DIANTSILAI TWKGRPDWTDYDFEQLPDRALTVFNPYNNRKSGRIVYGLRPGRSYOF 1330  
Db 910 QVAREITLHTWAPAG--EYTDFFELQYLSADEAPQLQNTKTEITLQGLRPHVHTF 967  
QY 1331 NVKTVSGDSWKT-----YSKPIFGSVRT---KPKIQNLHCRPQNSTAIACSW 1375  
Db 968 TVVVRSGSIQGTDFADVSVTLMRSSAPISASYQTLTAPPGKVDYFQSDVQPGVEVTEW 1027  
QY 1376 -IPP---DSDFDGVSIKCKMDTQVEFSRKLKESLLNMLVPHKRYLVSKVQSA- 1430  
Db 1028 SLEPAEOGRPIDYFRITQCNADDAADVSSFFPVNATQCKIDGLVPGNHYIFRIQKASAL 1087  
QY 1431 GWTSEVSDSTITMIDRPPPPPHIRVNEKDVLLSKSSINFTVNCWSFSDTNGVAKYFTV 1490  
Db 1088 GYGAEREHIQTPIL---APPVPEPSVTPLEVSRTSSSTIEISFROGYFSNAHGWRSYTI 1144  
QY 1491 VVREADGSDLEKPEQOQHPLPSYLEYRHNASIRVYQNTYFASKCAE--NP-----NSNSKSF 1544  
Db 1145 IIAEDVGKNASGLE---MPSQOD-----VQAY-TWMLPYQAIPEYNNPLTNSGRKS- 1192  
QY 1545 NIKLGAEMESIG-GKRDPTQOKFCGDLKPHYTRISRAFTQULFDEDLKFTPLYSDT-1603

Db 1193 --SLEAEHFTIGTANCXHQAGYCNGLRAGTTVRIKIRAPT---DED-----KFTDT 1240  
QY 1604 FFSLPITTESPLFGAIEGVSAGLFIQMLVAVVALLICROKVSHGRPRPSARLSIRDR 1663  
Db 1241 VYSSPITTERSDTVVAATVSA-VLLVAMVLVV---YCOHRCQ--LIRRAKSLARMQDE 1294  
QY 1664 PLSVHLNLGQGNKRTSCPIKINOPEGHFMKLOADSNYLLSKEYEELKXVGRNQSCDIAL 1723  
Db 1295 LAAL-----PEGYITPNRPVHVKDFSEHYRIMSADSRFSEEBEELKHVGRDQACSPAN 1349  
QY 1724 LPENRGNRYNNILPYDATRVKLSNVDDDDPCSDYINASYIIPGNFRREYIVTQGPLPGTK 1783  
Db 1350 LPCRPNRNFNIIIPYDHSRFLQPVDDDDDCSDYINANYMPOHNSPREFIVTQGPLHSTR 1409  
QY 1784 DDFWKMVEQNVHNVMTQCVKGRVKDHYWADQSDLYYGDILQLMLSESVLPEWTI 1843  
Db 1410 EEFWRMCWESNSRAIVMLTRCFEGREKCDQYWPVDMVAMFYGDIKVQLIITDTHVDWSI 1469  
QY 1844 REFKI---CGEEQIDAHRLIRHFHYTVWPDHGVFETTSLOIQTVRTVDYINRSFGAGPT 1900  
Db 1470 SEFWVRNRC-----ESRIMRHFHTTWPDPFGVPEPQSLVRFVRAFRDVI GTD--MRPI 1521  
QY 1901 VVHCSAGVGRGTFTIALDRILQLQSDSKSDVDIYGAVHDLRLHVRHVMVOTECQYVYLHOCV 1960  
Db 1522 IVHCSAGVGRSGTFTIALDRILQLHKSVDYDIFGLVFAMRKERVFMVOTECQYVCIHCL 1581  
QY 1961 RDVLARK 1968  
Db 1582 LAVLEGKE 1589

## RESULT 14

US-10-087-684-41  
; Sequence 41, Application US/10087684  
; Publication No. US20040029116A1  
; GENERAL INFORMATION:  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: MacDougall, John R.  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Stone, David J.  
; APPLICANT: Grosse, William M.  
; APPLICANT: Lepley, Denise M.  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Burgess, Catherine E.  
; APPLICANT: Casman, Stacie, J.  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Boldog, Ferenc L.  
; APPLICANT: Li, Li  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Tchernev, Velizar T.  
; APPLICANT: Vernet, Corine A.M.  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Gangoli, Esha A.  
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-214 CIP  
; CURRENT APPLICATION NUMBER: US/10/087,684  
; PRIORITY FILING DATE: 2003-03-10  
; PRIOR APPLICATION NUMBER: 60/253,834  
; PRIORITY FILING DATE: 2000-11-29  
; PRIOR APPLICATION NUMBER: 60/250,926  
; PRIORITY FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: 60/264,180  
; PRIORITY FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 60/274,194  
; PRIORITY FILING DATE: 2001-03-08

```
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: CuraSeqlist version 0.1
; SEQ ID NO 41
; LENGTH: 1767
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; US-10-087-684-41

Query Match      14.1%; Score 1475.5; DB 4; Length 1767;
Best Local Similarity 29.3%; Pred. No. 1.1e-84;
Matches 455; Conservative 257; Mismatches 662; Indels 180; Gaps 56;

Qy 518 KEFTFDLVPGRKY--MATVTSIGDGLKNSSVKGRTVPAQVTDLHVANQMTSGLFTNW 575
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 113 KDIKFSRALPGTEYNFMLYTTNSTHREQLTWVNTITAPDPPANLSVQLRSSKSAFITWR 172
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 576 TQAGDVEFYQVLLIHENVLIKESIS---SETSRYSFHSLSKGLSVVVVTVTSGGISS 632
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 173 PPGSGRVSFFIRVLGTLDPFERSYSLGNETLQLSAKETLPGGSYQVQAYSYYQKES 232
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 633 RQVVVEGRTPVSSVGVTVNNSGRNDYLSVSWL--VAPGDVDNYEVLTDLHSHDKVQVSLVI 690
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 233 VAYTSRNTTKPNTPGKFIWFRNETILLVLMQPPFPAGIYTHYRVSITPDDAIQSVLVY 292
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 691 AK-----SVRECSFSSLTGRLYTVTTTRSGKYENHSFS-----QERTVPDKVQGVSVS 740
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 293 EREGEPGPAQAQFKGLVPGREYNISVQTVS---EDETSSTPTTARYLTVPERVLNVTFD 349
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 741 NS-ARSDYLRVSW--VHATGDPDHYEVIKNNKFIQTKSPKSEN-ECYFVQ-----LVP 792
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 350 EAYTSSSFRVRWBPRTYSEFDAYQVMLSTSRIP---NVPRAAGDSVYFDYSDILEP 406
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 793 GRLYSVTVTKSGQVEANEQNGRTI-PEPVKDL--TLNRSTEDLHVTVSGA-NGDQVQ 848
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 407 GRTEVVVKTITADNVNWPASGEVTLRPRVRSLGGLDERSNA-LHISWEPATGRQDS 465
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 849 YEIQL--LFDNMKVFPFPP-----HLVNTATYRFTSLTPGRQYKILVLITSDVQOSAF 900
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 466 YRISYHEQTNASEVPAPFPAAESQITTNLTETLDSLLAGRRYLIAVQALSKGVASNAS 525
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 901 -IEGFTVPSA--VKNIIHSPNGATDSLTVNWTGGGDV-----DSYTVSAPRHSQKVDST 953
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 526 DITRYTPAAPLIQELRSIDQG-----LMLSWR-----SDVNSQDRYEVHYQRNGTR-BERT 577
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 954 IPKIVFEHTFRLRLEAGEQYQIMIASVSGSLKNQINNVGRTV-PASVQGVADNAYSYSYL 1012
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 578 MATNETSLTIHLHPGSGYEVKVHAI SHGVSEPHSYFQAVFPKPQNLTILQTVHTNL-V 636
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1013 IVSQKQAGVAERYDIIILLTENGILLNNTSP-----ATTQKHKFDLTPGKKYKIQIL 1066
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 637 VLHQAPEGSD-----FSEYVVRYTDSAPWQRISGLHENEARIKDMHYGERYLQVN 689
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1067 TVSGGLFSKEAQTRGRTVPAAVTDLRTENSTRHLSFRWTASEGELSWYNIFLNPPGNL 1126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 690 TVSGVSESPHLELNTVMPPOVSNVPLVDSRLNLTLEWPRPDGHVDFTLKWPTDE-- 747
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1127 QERQAVDPVQO-----SFSQNLQGRMYKM-VLVTHSGELSNESPIFGRTVP--- 1173
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 748 EDVFEKNVTVQLEDLSSPVRIPEDISPGHYRFEVQASNGIRSGTHLSITFTMLIQ 807
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1174 -----ASVSHLGRSGNRNTDLSLWFWNPASGD---PDFYELILYNPNKTKENWKDKLT 1225
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 808 SDVFIANAGHQGDETIT-----LSYTPTPADSTRFDIYRFSMGDPTIKDKEKLANDTER 863
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1226 EWRQGLVPGKRYVLVWVTHSGDLSN-KYTAESRTAPSPSLMSFADIANSTLAIWKGP 1284
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 864 KLSFSGLTGPKGLYNVTWTVSGGVALPVQRVYRHLPLPISDLKAIQVAAREITLHWTP 923
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1285 PDWTDYNDFELQWLPRDALTVFNPNYNNRKSGRIVYGLRPGRSYQFNKVTSGDSWKT-- 1342
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

; RESULT 15
; US-10-218-779-41
; Sequence 41, Application US/10218779
; Publication No. US20040029222A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit
; APPLICANT: MacDougall, John
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William
; APPLICANT: Alsobrook II, John
; APPLICANT: Lepley, Denise
; APPLICANT: Rieger, Daniel
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Spytak, Kimberly
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Patturajan, Meera
; APPLICANT: Shenoy, Suresh
; APPLICANT: Rastelli, Luca
```



APPLICANT: Tchernev, Velizar  
 APPLICANT: Vernet, Corine  
 APPLICANT: Zernusen, Bryan  
 APPLICANT: Malyankar, Uziel  
 APPLICANT: Guo, Xiaojia  
 APPLICANT: Miller, Charles  
 APPLICANT: Gangolli, Esha  
 TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
 FILE REFERENCE: 21402-214  
 CURRENT APPLICATION NUMBER: US/10/218,779  
 CURRENT FILING DATE: 2002-08-14  
 PRIOR APPLICATION NUMBER: 60/253,834  
 PRIOR FILING DATE: 2000-11-29  
 PRIOR APPLICATION NUMBER: 60/250,926  
 PRIOR FILING DATE: 2000-11-30  
 PRIOR APPLICATION NUMBER: 60/264,180  
 PRIOR FILING DATE: 2001-01-25  
 PRIOR APPLICATION NUMBER: 60/313,656  
 PRIOR FILING DATE: 2001-08-20  
 PRIOR APPLICATION NUMBER: 60/327,456  
 PRIOR FILING DATE: 2001-10-05  
 NUMBER OF SEQ ID NOS: 216  
 SOFTWARE: PatentIn ver. 2.1  
 SEQ ID NO 41  
 LENGTH: 1767  
 TYPE: PRT  
 ORGANISM: Drosophila melanogaster  
 US-10-218-779-41

Query Match 14.1%; Score 1475.5; DB 4; Length 1767;  
 Best Local Similarity 29.3%; Pred. No. 1.1e-84;  
 Matches 455; Conservative 257; Mismatches 662; Indels 180; Gaps 56;

Qy	518	KEFTFDLVPGRKY--MATVTSISGDLKNSSVKGRTVPAQVTDLHVANQMTSSLEFNW	575
Db	113	KDIFSRALPCTEYNFWLYTNTSHREQLTWTNITAPDPPANLSVQLRSKSAFITWR	172
Qy	576	TOAGDVEFYQVLLIHENNVTKNISIS---SETGRYSFHSLSKSGSLYSVVVTVTSGGISS	632
Db	173	PPSGRYSGPFRIRVLGLTDLFPERSYSLGNETLQLSAKELTPGGSYQVQAYSYYQKES	232
Qy	633	QOVVVEGRTVPSSVSGVTNNSGRNDYLSVGL--VAPGDVDNVEVTLSDHGKVVQSIVI	690
Db	233	VAYTSRNPFTKPNTPGKFIWFRNETLLVLWQPPFAGIVTHYRVSITPDDAIQSVLYV	292
Qy	691	AK-----SVRECSFSSLTTPGRLYTVITTRSGKYENHSFS-----QERTVPDKVQGSVS	740
Db	293	EREGETPGPAQAFAKGLVPGREYNLSVQTVS---EDETSSVPTTARYLITVPERVLNFTD	349
Qy	741	NS-ARSDVLRVSW--VHATGDFHYEVTIKNNFIQTKSIPIKSEN-ECVFPVQ-----LVP	792
Db	350	EAYTSSSFRVRWEPPTYSBFDAYQVWMLSTRIF---NVPRAANGDSVYFDYSILEP	406
Qy	793	GRLSVTVTTKSGQYEANEQNGRTI-PEPVKDL--TLNRSTEDLHVWMSGA-NGVDVQ	848
Db	407	GRTEVWVKTIADNNVNSMPASGEVTLRPRVRSLSGLGFLDRSNA-LHISWEPAETGRQDS	465
Qy	849	YEIQI--LFNDMKVFPPP-----HLVNTATEYFTSLTPGRQYKILVLTISGDVQOSAF	900
Db	466	YRISYHEQTNASEVPAPPVAAESQITNLTEYILDSLAGRRVLIIVQALSKGVSANAS	525
Qy	901	-IEGFTVESA--VRNIHISPGATDSLTVNWTTPGGDV-----DSYTSVAFRHSOKVDSQT	953
Db	526	DITRYTPAAPLIQELRSIDQ-----LMSWR---SDVNSRQDRSYEHVYQNGTR-EERT	577
Qy	954	IPKHVFETFRLEAGEYQIMIASVSGSLKNQINVGRTV-PASVQVIADNAYSYSVL	1012
Db	578	MATNETSLTIHLPGSGYVKVHAISHGVSEPHSYFQAVFPKPPQNLTQTVHTNL-V	636
Qy	1013	IVSQKAGVAERYDILLTENGILLRNTSEP-----ATTQKHFDLTGPKYKIQIL	1066
Db	637	VLHWQAPGSD-----FSEYVRYRTDASFPWRISGLHENEARIKDMHYGERYLQVN	689

Search completed: March 10, 2006, 19:06:41  
 Job time : 203.859 secs

Qy	1067	TVSGGLFSKEAQTEGRTVPAAVTDLRLITENSTRHLSFRWTASEGELSWMYIFLYNPDGNL	1126
Db	690	TVSPGVESPHLELNVTMPQPVNSVNVPLVDLWPRPDGHVDFYTLKWPFTDB--	747
Qy	1127	QERAQVDPLVQ-----SFSFQNLQGRMYKM-VIVTHSGLSNESIFGRTVP---	1173
Db	748	EDRVEFKNVTOLEDLSSPSVRIPIEDLSPCQYRFEVQASSNGIRSGTTHLSTRIMPLIQ	807
Qy	1174	-----ASVSHLRGSRNRTTDSLWNWSPASGD-----FDYELILYNPNTKKNWKDKDLT	1225
Db	808	SDVFIANAGHEQODETIT-----LSYTFTPADSTRFDIYRFSMGDPTIKOKEKLANDTER	863
Qy	1226	EMRFQGLVPGRKVYLWVTHSGDLSN-KVTAESTAPSPSLMSFADIANSTSLAITWKGK	1284
Db	864	KLFSGLTUPGKLYNVTVTVSGGVASLPVORVYRLHPLIPISDLKAIQVAREBITLHWTP	923
Qy	1285	PDWTDYNDFELQWLPRDALTVFNPNRNRKSEGRIVYGLRFGRSYQFNKTVSGDSWKT--	1342
Db	924	AG--EYTDLELOYLSDAEAPQLLQNTVTKTEITLQGLRPVHNVTFTVVRSGSISQGTDF	981
Qy	1343	-----YSKPIFGSVRT---KPKIQNLHCRPONSTAIACSW-IPP---DSDPFGY	1385
Db	982	ADSVSVTLMRSSAPISASYQTLTAPPGKVDFYQPSDVQPGGEVTFEWSLEPAEQHGPIDYF	1041
Qy	1386	SIECRKMDTOEVEFSRKLKEKESLLNTMLVPHKRYLSVKVQSA-GMTSEVVEDSTITM	1444
Db	1042	RITCONADDAADSVSYEPFNATQGTIDGLVPGNHYIFRIQAKSALGAGAREHIQNTMPI	1101
Qy	1445	IDRPPPPPHIRVNEKDVLSKSFNTVNCSPFSTNGAVKYFTVVVVRREADGSDDELKPE	1504
Db	1102	L---APPVPSVTPLEVSRTSSTIEISFRQGYFSAHGMVRSYTIILAEVDGK-----	1152
Qy	1505	QQHPLPSYLEYRNHASIRVQNTNYFASKAE--NP-----NSNSKSFNIKUGAEMESLG-G	1557
Db	1153	-----IASGLEMPSNQDVQAY--TWMLPYQAIPEYNPFLTNSGRKS---SLEAEHLTIGTA	1204
Qy	1558	KRPTQKFCGDLKPHYATYRISITRAFTQLFDEDLKEFTKPLYSDTFSPILPITTESEPLF	1617
Db	1205	NCDKHQAGYCNGLPAGTYYRIKIRFT---DED-----KFTDTVYSSPITTERSDTV	1254
Qy	1618	GATEGVSAFLIGMLVAVVALLICROKVSNGRERPSARLSIRDRPLSVHLNLGQKGNR	1677
Db	1255	IVAATVSA-VLLVAVLVVV---YQHRQ---LIRASKLARMQDELAAL-----PEGVI	1303
Qy	1678	KTSCPIKINOFEGHFMKQADSNYLLSKEVEELKDVGRNQSCDIALPENRGNKRYNNIL	1737
Db	1304	TENRPHVKDFSEHYRIMSADSDFRFSEPEELKHVGRDQACSPANLPCNRPKNFTNIL	1363
Qy	1738	PYDATRVKLSNVDDPCSDYINASYIPGNFRREYIVTQGLPCTKDDFKWVMEQVHN	1797
Db	1364	PDHSREKLPQVDDDDGSDYINANMPGHNSPREFIVTQGFPHSTRBEFWMWCWESRA	1423
Qy	1798	IVMVTQCEKGRVKCDHYWADODSLYGDLLIQLMESVLPETIREFKI---CGEEQL	1854
Db	1424	IVMLTRCFEKGREKCDQYWPVDRVAMFYGDIKVQLIIDTHYDWSISEFVSRNC-----	1478
Qy	1855	DAHLRIHFHYTWPDPHGVPEPPTQSLIQFVRTVDYINRSPAGTVPVHCSAGVGRGTTF	1914
Db	1479	-ESRIMRHFHTTWPDPGVPEPPLSLRVFVRAFRDVIQTD--MRPIIVHCSAGVGRSGTF	1535
Qy	1915	IALDRILQQLDSDKSDIYGAVHDLRLHRVHMVQTECYVYLHQCVDRDLARK	1968
Db	1536	IALDRILQIHKSDYDIFGIVFAMRKERVFMVQTEQQYVCIIHOCCLAVLEGKE	1589

Page Blank (uspto)

Page Blank (uspto)

November 2005

Published\_Applications\_Nucleic Acid and Published\_Applications\_Amino Acid database searches now generate two sets of results each. The Published\_Applications\_databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications\_New databases; older published applications make up the Published\_Applications\_Main databases.

Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New). Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.rapbm** (Published\_Applications\_AA\_Main) and **.rapbn** (Published\_Applications\_AA\_New).

**This Page Blank (uspto)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2006, 19:00:41 ; Search time 21.3095 Seconds.  
(without alignments)  
2608.526 Million cell updates/sec

Title: US-10-633-742-2

Perfect score: 10483

Sequence: 1 MLSHGAGLWITLSLQGTG.....PPIYVNVPEYHRDPVYSRH 1997

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 161667 seqs, 27834885 residues

Total number of hits satisfying chosen parameters: 161667

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US03\_NEW\_PUB pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1242	11.8	1337	US-11-112-304A-33	Sequence 33, Appl
2	1232	11.8	1178	US-10-995-561-851	Sequence 851, Appl
3	1204	11.5	1705	US-11-143-984A-37	Sequence 37, Appl
4	1198	11.4	1711	US-11-143-984A-38	Sequence 38, Appl
5	913.5	8.7	1188	US-11-143-984A-27	Sequence 27, Appl
6	832	7.9	1907	US-11-000-463-250	Sequence 250, Appl
7	826	7.9	1897	US-10-821-234-1635	Sequence 1635, Ap
8	780.5	7.4	405	US-11-143-984A-28	Sequence 28, Appl
9	756.5	7.2	1254	US-10-528-031-47	Sequence 47, Appl
10	718.5	6.9	1452	US-10-821-234-1102	Sequence 1102, Ap
11	691.5	6.6	1463	US-11-080-991-22	Sequence 22, Appl
12	687.5	6.6	2314	US-11-097-728-2	Sequence 2, Appl
13	687.5	6.6	2353	US-11-097-728-6	Sequence 6, Appl
14	675.5	6.4	2477	US-11-193-561-15	Sequence 15, Appl
15	675.5	6.4	2477	US-11-193-771-15	Sequence 15, Appl
16	675.5	6.4	2477	US-11-193-789-15	Sequence 15, Appl
17	675.5	6.4	2477	US-11-193-806-15	Sequence 15, Appl
18	675.5	6.4	2477	US-11-193-857-15	Sequence 15, Appl
19	672	6.4	1445	US-11-169-041-181	Sequence 181, Appl
20	671	6.4	2223	US-11-193-561-2	Sequence 2, Appl
21	671	6.4	2223	US-11-193-771-2	Sequence 2, Appl
22	671	6.4	2223	US-11-193-789-2	Sequence 2, Appl
23	671	6.4	2223	US-11-193-806-2	Sequence 2, Appl
24	671	6.4	2223	US-11-193-857-2	Sequence 2, Appl
25	662.5	6.3	2421	US-11-193-561-17	Sequence 17, Appl

Sequence 17, Appl  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 626, App  
Sequence 722, App  
Sequence 623, App  
Sequence 627, App  
Sequence 19, Appl  
Sequence 19, Appl  
Sequence 19, Appl  
Sequence 19, Appl  
Sequence 1545, Ap  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 21, Appl  
Sequence 38, Appl

26 662.5 6.3 2421 7 US-11-193-771-17  
27 662.5 6.3 2421 7 US-11-193-789-17  
28 662.5 6.3 2421 7 US-11-193-806-17  
29 662.5 6.3 2421 7 US-11-193-857-17  
30 608 5.8 2386 6 US-10-995-561-626  
31 607.5 5.8 647 7 US-11-000-463-722  
32 601.5 5.7 2355 6 US-10-995-561-623  
33 601.5 5.7 2355 6 US-10-995-561-627  
34 601.5 5.7 2355 7 US-11-193-561-19  
35 601.5 5.7 2355 7 US-11-193-771-19  
36 601.5 5.7 2355 7 US-11-193-789-19  
37 601.5 5.7 2355 7 US-11-193-806-19  
38 601.5 5.7 2355 7 US-11-193-857-19  
39 601.5 5.7 2384 6 US-10-821-234-1545  
40 595 5.7 2330 7 US-11-193-561-21  
41 595 5.7 2330 7 US-11-193-771-21  
42 595 5.7 2330 7 US-11-193-789-21  
43 595 5.7 2330 7 US-11-193-806-21  
44 595 5.7 2330 7 US-11-193-857-21  
45 585.5 5.6 2217 7 US-11-193-561-38

#### ALIGNMENTS

##### RESULT 1

US-11-112-304A-33

; Sequence 33, Application US/11112304A

; Publication No. US20060002931A1

; GENERAL INFORMATION:

; APPLICANT: AMGEN, INC.

; APPLICANT: Smothers, James

; APPLICANT: Fanslow III, William C.

; APPLICANT: Kariv, Revital

; TITLE OF INVENTION: ANTIBODIES OF ANGIOGENESIS INHIBITING DOMAINS OF CD148

; FILE REFERENCE: 3447

; CURRENT APPLICATION NUMBER: US/11/112.304A

; CURRENT FILING DATE: 2005-04-22

; PRIOR APPLICATION NUMBER: US 60/565,158

; PRIOR FILING DATE: 2004-04-23

; PRIOR APPLICATION NUMBER: US 60/564,885

; PRIOR FILING DATE: 2004-04-23

; PRIOR APPLICATION NUMBER: US 60/571,566

; PRIOR FILING DATE: 2004-05-14

; PRIOR APPLICATION NUMBER: US 60/585,686

; PRIOR FILING DATE: 2004-07-06

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 33

; LENGTH: 1337

; TYPE: PRT

; ORGANISM: Homo sapiens

US-11-112-304A-33

Query Match 11.8%; Score 1242; DB 7; Length 1337;

Best Local Similarity 29.6%; Pred. No. 7.2e-75;

Matches 397; Conservative 190; Mismatches 503; Indels 251; Gaps 53;

Qy 759 FDHYEVTKNNFIQTKSP-----KSENE-----CVFVLVPGELY 796  
Db 123 FDIKAVSISPTNVILTWKSNDAASEYKYVYVGHMENEKTIIVVHQPWCNITGLRPATSY 182  
Qy 797 SVTVTKSGGYANEQ-GNGRTI-----PEPVKDLTUNRSTEDLHVTWSGANG----- 844  
Db 183 VFSITPGIG-----NETWGDPRVIVKITEPIPVSDLRVALTGVKKAALSWNGNGTASCRV 238  
Qy 845 ---DVDQYEIQLLFNDMKVFPFHLVNTATEYFTSLTPGRQYKILVLTISGDVQQAFAI 901  
Db 239 LLESIGSHE-ELTQDSRL-----QVNISDLKPGVQY-----NINPYLQSNKT 280  
Qy 902 EGGTVPSAVKNHISIPNGATDSLTWNTPGGGDVDSVTVSFAFRHSQKVDSTQITPKH----- 957

Db 281 KG-----DPLG-----TEGLDASN-----TERRSAGSPTAPVHDESL 313  
Qy 958 -----VFEHTFHRLEAGEQYQIMIAS--VSGSLKNQINVVGRVTPASVQGV 1002  
Db 314 VGPVDPSSGQSRDTEVLVGLGPGTRNATVYQAANGTEGQQAIEFRNATQVDFVT 373  
Qy 1003 ADNAYSSYLIVSWQKAAGVAE---RYDILLTENGILLRNTSPPATTQKHKFDLTPGK 1059  
Db 374 AVN-ISATSLTLIWKVSDNESSNVTYKIHVAGETDSSNLNVSEPR-----VIPGL 424  
Qy 1060 K-----YKIQILTVSGGLFSKAQTEGRTVPAAVTDLRLITENSTHLSFRWTASGELSWY 1115  
Db 425 RSSTFYNTICPVLDIGTEGFLQVHTPPVPVDSFRVTVTSTTEIGLAWSSHDAESFQM 484  
Qy 1116 NIFLYNPDNLQERAQVDPLVQSFSQNLQGRMYKMVIVTH--SGELSNESFPGRTPV 1173  
Db 485 HI-----TQEGAGNSRVEI--TTNQSIILGGLPGTKYCFEIVPKGNGTEGASRTVCNRTVP 540  
Qy 1174 ASVSHLRGSNRTTDSLWFMNPSAGDPDF-YELILYNPNGTKKENMKDKDLTEWRFGQL 1232  
Db 541 SAVFDIHVVYVTTTE-MWLDWKS PDGASEYVYVHLVIESKHGNSHTSTYDKAIT--LQGL 596  
Qy 1233 VPGKRYLVWV---VTHSGDLNKNVTAESRTAPSPSLMSPADIANTSIAITWKGPDPWTD 1289  
Db 597 IPGLTYNITISPEVDHVMGDPNSTAQYTR-----PSNVSNIDVSTNTTAATL-----SWQN 647  
Qy 1290 YNDFELOWLPDRLALTVPFNPNRKSSEGRIVYG-----LRPGRSYQFNKVTSGDS 1339  
Db 648 FDDASPTV---SYCLLIEKAGNSNATQVTDIGITATVTELIPGSSYTVEIFAQVGDG 704  
Qy 1340 WKYSKPIFGSVRTKPKQIONLHCR-PQNSTAIACSWIPDSDPDGYSIE-----CR 1390  
Db 705 IKSL-EPGRKSFCTDPASMASFDCVWPKEPALVLKWTCPGANAGFELEVSSGAWNAT 763  
Qy 1391 KMTQVEFFSKLEKESLNLMMVLPHKRYLVLSIKVQAGMTSEVEDSTITMIDRPPP 1450  
Db 764 HLESCSSENGTEYRTEVTVLNF-----STSNISITTVSCGKMAAPTRNTCTGITDPPP 818  
Qy 1451 P--PPIRVNKEVDLISKSSNFTVNGSWFSDTNGAVKFTVVVREADGSDDELAKPEOHP 1508  
Db 819 PDGSPNI-----TSVSHNSVK--VKFSGFASGPIKAVAVILTTG-----EAGHP 862  
Qy 1509 LPSYLEYRHNASIRVYQTNV--FASKCAENPNSNKS-----FNIKLGAEMESLGGKRDPT 1562  
Db 863 SADVLKYTYDDFKKGASTVYTYLIRTEEKGSRQSLSEVLKYEIDVGNESITLG-----916  
Qy 1563 QOKFCGDLKPHATYAIRISIRAFQ-----LFD--EDUKETKPLYSDTFPSLPIIT 1611  
Db 917 ---YNGKLEPLGYSRACVAGFTNITFHPQNKGLIDGAESYVSPSR--YSDA-VSLP---967  
Qy 1612 ESEP--LFGATEGVSAGIFLGLMLVAVVALLICROKYSHGREPERSARLSIRDRPLSVHL 1669  
Db 968 -QDPGVICGAVFGCIFGALVI---VTVGGFIFWRKK---RKDAKNVESFSQIKP-----1015  
Qy 1670 NLGQGNRKTSCPIKINQFEGHFMKLOADSNYLLSKEYEELKVGYNQSCDIALLPENRG 1729  
Db 1016 -----KSKLIRVENFEAYFKKQADSNCGFAEEVEDLKLVGISOPKYAAELANRG 1067  
Qy 1730 KQRNNILPYDATRVKLSNVDDDDPCSDYINASYTPGNFRREYIVTQGPPLPGTKDDFWKM 1789  
Db 1068 KNRNNVLPYDISRVKLS-VQTHSTDDYINANYPGYSHSKKDFATQGPLNTLKDFFWRM 1126  
Qy 1790 VWEONVNIVMVTOCVKGRVKCDHYWPAODSLYGLDILQLMSESVLPFWTIREPKIC 1849  
Db 1127 VWEKNVAILMLTKCVSGRTRKCEYWPSSQ-AQYGDITVAMTSEIVLPEWTIRDFTVK 1185  
Qy 1850 GEEOLDARLIRHPHYTVDPHGVPETTTQSILIQFVRTVRYDINRSPGAGPTVHVCSAGV 1909  
Db 1186 NIQTSSEHPL-RQHFHTSPDHGVPDITDILLINFRYLVRDYMKQSPPEPILVHCSAGV 1244  
Qy 1910 RTGTFIALDRILQOLDSKSDVIDYIGAVHDLRLHRVHMVQTECQYVYHQCVRDVLARKL 1969  
Db 1245 RTGTFFIADRUIQIENENTVDVYIGIYVYDLRMHRLPLMVQTEBDQYVFLNQCVLDIVRSQK 1304

Qy 1970 RS-----EQENPLFPPIYENVNP 1986  
Db 1305 SKVDLIYQNTAMTIYENLAP 1325  
RESULT 2  
US-10-995-561-851  
; Sequence 851, Application US/10995561  
; Publication No. US20050272054A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
; TITLE OF INVENTION: DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001559  
; CURRENT APPLICATION NUMBER: US/10/995,561  
; CURRENT FILING DATE: 2004-11-24  
; NUMBER OF SEQ ID NOS: 85702  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 851  
; LENGTH: 1178  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-995-561-851

Query Match 11.8%; Score 1232; DB 6; Length 1178;  
Best Local Similarity 29.7%; Pred. No. 2.8e-74;  
Matches 390; Conservative 182; Mismatches 485; Indels 256; Gaps 53;

Qy 782 ENE-----CVFVQLVPGRLYSVTVTTKSGQVEANEQ-GNGRTI-----PEPVKDL 825  
Db 2 ENEXTITVVHQPWCNITGLRATPSYVFSITPGIG---NETWGDPRVIKVTIPEIPVSDL 57  
Qy 826 TLRNRSTEDLHVTWSGANG-----DVDQVEIQLLFNDMKVFPFPHLVNTATEYRFT 876  
Db 58 RVALTGVKAAKLSWNGSNGTASCRLLESIGSHE--ELTQDSRL-----QVNIS 104  
Qy 877 SLTPGROYKILVLTISGDVQOQSAFIEGFTVPSAVKNHIHPNGAT--DSLTVNWTGGGD 934  
Db 105 GLKPGVOYNI-----NPLYLPSNKTGKDDL--GTEGLD 136  
Qy 935 VDSYTVSAFRHSQKVDSTQIPKH-----VFETHFRLEAGEQYQIMIA 977  
Db 137 ASN-----TERSRAGSPTAPVHDESLVGPVDPSSGQSQDTEVLVGLGPGTRYNATVY 190  
Qy 978 S--VSGSLKNQINVVGRVTPASVQGVADNAYSSYLIVSWQKAAGVAE---RYDILLT 1032  
Db 191 SOAANGTEGQQAIEFRNATQVDFVTAVN-1SATSLTIWKVSDNESSNVTYKIHVAG 249  
Qy 1033 ENGILLRNTSEPTATTKQHKFEDLTPGKK---YKIQILTVSGGLFSKAQTEGRTVPAAV 1088  
Db 250 ETSSNLNVSEPR-----VIPGURSTFTNITVCPVLGIEGTPGFLQVHTPPVPV 301  
Qy 1089 TDLRITENSTHLSFRWTASGELSNVTNIFLYNPDGNLQERAQVDPLVQSFSQNLQGR 1148  
Db 302 SDFRVTVTSITTEIGLAWSSHDAESFQWHI---TQEGAGNSRVEI--TTNQSIILGGLPPT 357  
Qy 1149 MYKRVIVTH--SGELSNESIFORTVTPASVSHLRGNSRNTTDSLWFMNWSPASGDPDF-YE 1205  
Db 358 KYCFEIVPKGPNTEGASRTVCNRTVPSAVFDIHVVYVTTTE-MWLDWKS PDGASEYVYH 416  
Qy 1206 LILYNPNGTKKENMKDKDLTEWRFGQLVPGKRYLVWV---VTHSGDLNKNVTAESRTAPS 1262  
Db 417 LVIESKHGNSHTSTYDKAIT--LQGLIPGTLNITISPEVDHVMGDPNSTAQYTR-----469  
Qy 1263 PPSLMSFADIANTSIAITWKGPDPDITDYNDFELQWLPRDALTVFNPYNNRKSSEGRIVYG- 1321  
Db 470 -PSNVSNIDVSTNTTAATL---SWQNFDASPTY---SYCLLIEKAGNSNATQVVTDI 521  
Qy 1322 -----LRPGRSYQFNKVTSGDSWKYTKSPIFGSVRTKPKQIONLHCR-PQNSTAI 1371  
Db 522 GITDATVTELIPGSSYTVEIFAQVGDGIKSL-BFGRKSFCTDPASMASFDCVVPKBPAL 580

Qy 1372 ACSWIPDSDFDGYSIE-----CRMDTQVEFSRKLKESLLNIMMLVPHKRYLV 1423  
Db 591 VLKWTCPGANAGFELEVSSGAMNATHLESCESENGETEYTEVYTLNF-----STSYNI 635  
Qy 1424 SIKVQAGMTSEVVEDSTITWIDRPPPP--PPHVRVNEKDVLISSKSNFTVNCWSFSDT 1481  
Db 636 SITVSCCKMAAPTCTTGTIDPPPDGSPNI-----TSVSHNSVK--VKSGGFAS 687  
Qy 1482 NGAVKYFTVVVREADGSDDELPEQOHPSPSYLEYVRHNASIRYQNY--FASKCAENPNS 1539  
Db 688 HGPIKAYAVILTTG-----EAGHPGADVLKTYEDFKGASDTYVYLIRTEEKGRS 739  
Qy 1540 NSKS-----FNIKLGAEMLGCKRDPTQCKDCDPLKETHAYRISIRAFQ----- 1586  
Db 740 QSLSEVLKYEIDVGNESITLG-----YNGKLEPLGSTRACVAGFTNITHPQWKG 790  
Qy 1587 LPD--EDLKEFTKPLYSPTTFSLPITTESEP--LFGAIEGVSAGLFLTGLMLVAVVALLIC 1642  
Db 791 LIDGAEVYSFSP--YSDA-VSLP-----QDPQVICGAVFGCIFGALVI---VTVGGFIFW 840  
Qy 1643 RQKVSHGRERPSARLSIRDRPLSVHLNLGQGNKRTSCPIK-----INOPEGHFMKLOA 1697  
Db 841 RKK-----RKD-----AKNEVSPSQIKSKLIRVENFEAVFKQQA 876  
Qy 1698 DSNVLLSKEYELKDVGNQSCDIALPENRGKRYNNILPYDATRVKLSNVDDDDPCSDY 1757  
Db 877 DSNCGFABEYEDLKLVGISQPKYAAELAEKRGKRYNNVLPYDISRVKLS-VQTHSTDDY 935  
Qy 1758 INASYIPGNFRREVIVTQGLPGTKDDFKWVWQEQNVHNTVWVTCVKEKGRKCDHWYP 1817  
Db 936 INANTMPGYHKKDFIATQGLPNTLKDFWRWWEKKNVYALIMLKCVQEQRTKCEEYWP 995  
Qy 1818 ADQPSLYYGDILQMLSESVLPEWTIREFKICGEEQLDAHLRIRHFHYVWPDHGVPETT 1877  
Db 996 SKQ-AQDYGDITVAMTSEIULPEWTIRDTVKNIQTSSHEL-RQFHFTSPEDHGVPEDTT 1053  
Qy 1878 QSLQFVTRVDYINRSAGPTVWHCSAGVGRGTGTFALDRILQOLDKSDSDIYGAVH 1937  
Db 1054 DLLINFRYLVDYKQSPSPILVHCSAGVGRGTGTFALDRILYQIENENTVDYGVIV 1113  
Qy 1938 DLRHVRHVMTQEQYVVLHOCVRDLRARKLS-----EQENPLFPPIYENVP 1986  
Db 1114 DLRHVRPLMVQTEQYVFLNQCVDLIVRSQDKVDLLIYQNTAMTIYENLAP 1166

## RESULT 3

US-11-143-984A-37  
; Sequence 37, Application US/11143984A  
; Publication No. US20060014180A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL HUMAN PHOSPHATASES  
; FILE REFERENCE: D0072 DIV1  
; CURRENT APPLICATION NUMBER: US/11/143,984A  
; PRIOR FILING DATE: 2005-06-02  
; PRIOR FILING DATE: 2000-12-20  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR FILING DATE: 2001-05-01  
; PRIOR FILING DATE: 2001-05-01  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR FILING DATE: 2001-06-25  
; NUMBER OF SEQ ID NOS: 208  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 37  
; LENGTH: 1705  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-11-143-984A-37

Query Match 11.5%; Score 1204; DB 7; Length 1705;  
Best Local Similarity 25.1%; Pred. No. 3.7e-72;  
Matches 411; Conservative 275; Mismatches 671; Indels 282; Gaps 52;  
Qy 454 NLKNSERCQGRTRVPLAV-LQLRVKHAHNETSLIMWQTPVAEWEKYIISLADRLD----- 506  
Db 25 SLDSGPRQGGPPLSVNVSRRGK---PTSLFLSWAAEPCGFDYALCLRAMNLGSPFEG 81  
Qy 507 --LLIHKSLSKDAKEFTFTDLVPGKRYMATVTSISGLDN-SSSVKGRTPAQVTDLHVA 563  
Db 82 QQLQAHNTES-----SFEHGLVPGSRVQLELTVLRPCWQNVITILTARTAPTAVRGLQLH 137  
Qy 564 NOGWTSSIFTNWTQACQDVEFYQVLLAH--ENVVIKNESISSETSRYSFHLKSGSLYSV 621  
Db 138 STGSPASLESWSASDQSDSYQLLLYHPESHHTLACNVSVSPDTLSYNFGLLLPGSQVYL 197  
Qy 622 VVTVSGIGISRRQVVVSGRTVPSSVSGTVVNNSGRNDYLSVSLVAPGDVDVNYETVLSH- 680  
Db 198 EVITWASLHAKTSILOH-TEPVPPDHLTLRALGTSS-LQAFWNSSEG-----ATWHL 249  
Qy 681 -----DGKVVQSILVIAKSVRECSFSLTPGRLYTVTTTTRSGKYENHSFQERTVPKV 734  
Db 250 ILTDLLEGTNL-TKVVRQGISHTFLRLSPGTYPQLKICAAAG-----PHQI 295  
Qy 735 QGVSVNSARSXY-----LRVSWHATGDFDHYEVTIKNKNFIQTKSI PKSE 782  
Db 296 WGNATEWTPSYSPSDLVLTPLMNLWASWKAGQAGRDGY--VLKLSGVENTTTLGPEE 353  
Qy 783 NECVFLQVLPGRLYSVTVTKSGOYEANEQCN-----GRTIPEPVKDLTLNRSTEDL 835  
Db 354 CNAVFPGLPFGHYTLGLRVLAGPYDAWEGSIWLAESAARPMVEVPGARLML-----EGL 408  
Qy 836 HVT--WSGANGVDQVYEQILLFNDMKVPFPPHVLVNTATEYRFTSLTPGQYKILVLTISGD 894  
Db 409 EATKQPCRRALLYSDAPGLLGNI SV-----SSGATHVTCGLVPGAHYRVDIASSMGD 462  
Qy 895 VQOSAFTEGFTVPSAVKNIHISPNAGTDSLTVNWTGCGDVSYSVSAFRH-SQKV--DS 951  
Db 463 ITQS--LTGYTSPLPPOSLEIISRNPSDLTIGWAPAPQOMEGYKVTWHQDGSQRSPDL 520  
Qy 952 QTI PKHVFETHFRLAGEOYQIMIASVSGSLKQINVV--GRTPVPSVQGVIANAYSSY 1010  
Db 521 VDLGPDISLTLKSLVPGSCYTVSAWAGNLSDDSKIHSCITRAPPTNLISLGAHQPA 580  
Qy 1011 SLIVSWQAAAGVARYDILLTENGILLRNTSEPAATKQH-KFBDLTPGKIKYQILTVS 1069  
Db 581 TLRASWCHPPGGRDAFQLRLVRLRPLTLESEKILSQAQNFSAQLPAGYEFQVQLSTLW 640  
Qy 1070 GGLFSKEAQTEGRTVPAAVTDLRITENSTHLSFRWTASEGELSWYNIFLNPDGNLOER 1129  
Db 641 GSESGSANTTGTWTPPSAPTLLVNVTSBAPTQLHVSWHAAGDRSSYQVTLTYQ-ESTRTAT 699  
Qy 1130 AQVDPLVQSPFQNLLOGRMVKMIVITHSGEL-SNESIFGRTVTPASVSHLRGSRNRTD 1188  
Db 700 SIVGPKADSTFWGLTGTKYKVEAISWAGPLYTAAANVSANWTPLTPNELLASQAQSA 759  
Qy 1189 SLWFNWPASGDFDFEYLILYNPNGTCKENWKDLEWRPQGLVPGRKYYVLVWVTHSGD 1248  
Db 760 VVNLAW-----PSGLQGTCHQLSD-----AGH 784  
Qy 1249 LSNKVTAESRTAPSPSLMSFADIANSTLAITWKGPDPDWTNDYDFELQWLPRDALTVPNP 1308  
Db 785 LS-----WQPLS-----LQODLLMLRN- 802  
Qy 1309 YNNRKSGRIVYGLRPGRSYQFNVTYSGDSWKYTSKPIFGSVRTKDKIQLNLCRPNOS 1368  
Db 803 -----LIPGHTVSLSKCRAG-PLQASTHPLVLSV--EPGPVEDVFCQPE-A 845  
Qy 1369 TAIACSWIPDSDFDGYSIECKRM-DTQEVEFSRKLKESLLNIMMLVPHKRYLVSKV 1427  
Db 846 TYLSLNTMPTGADVAVCLVEQVLVPGSAHFVQVNTSEADALLPLNLTPTTYSRLSLTV 905

```
Qy 1428 QSAG-WTSEVVEDSTITMIDRPPPP-----PPHVRNEKDVLSKSSINFTVNCWSFSDTN 1482
Db 906 LGGNRQWSRAVTLVCTTSAEVAHPPELAEAQVQLG-----TGMGVTVTRGMFGKDD 957
Qy 1483 GAVKYFTVVREAGSDELKPEQHPPLPSYLE--YRHNASIRVQTYFASKCAENPSN 1540
Db 958 GOIQWYGIIA--TINMTLAQPSQEAINHNTWYDHYRGHDS---YLALLFPNPFYPEPWAV 1012
Qy 1541 SKSNIKLAGWESLGGKRDPTQCKDGPGLKPHPTAYRISIRATQTFQDFED---LKEFT 1596
Db 1013 PRSTVTPVGTB-----DCDNTQB-ICNGHLKPGQVRFSAARSLSSPETIILAFSAFS 1065
Qy 1597 KPLYSDTFPSPIITTESEPLFAGIEGVSAGFLIGMLVAVVALLICRQVSHGRERPSAR 1656
Db 1066 EPOASISLVAMPPLTWMTGVVGGCI-----IIVCAVLCLLCRRLLKGRPRSEKNG- 1113
Qy 1657 LSIRDRPLSVHL---NLGQKGNKRTCPKINQFEGHFMKLQADSNYLLSKEYEELKDV 1713
Db 1114 -----FSQELMPYNLW-----RTHRPISHSFROSYEAKSARAHQAFFQEFELKEV 1160
Qy 1714 GRNOSCDIALLPENRGNRYNNILPYDATRVKLSNVDDDDPCSDYINASYIPGNFRREYI 1773
Db 1161 GKQDRLAEAHANITKORYPHVLPYDHSRVLRLQLSGEPHSDYINANFIPGYSHPOEII 1220
Qy 1774 VTQGLPGTKDDFKWVMEQNVHNIVMVTCQVGRVKCDHYWPADQDSLYYGLIQLML 1833
Db 1221 ATQGLPKTVEDEFWRLVWEQVHVIMLTVMGMENGRVLCHEHYWPNSTPVTGHITHTLL 1280
Qy 1834 SESVLPDWTITREFKI-CGEQDLAHLRIHRYHTVDPDGHVPEPTQSLIOFVRTVRDYIN 1892
Db 1281 ABESEDEWTRREFOLHGAEQ--KQRRVKVQLQFTTWPDSVPEAPSSLLAFVELVQBEVK 1338
Qy 1893 RSPCAGPTVHCASGAGRTGTFIALDRLIQLQDSKSDYIYGAVHDLRLHRVHMVQTECO 1952
Db 1339 ATQKGPLVHCASGAGRTGTFVALLPAVROLEBEEQVVDVNTVYILRLHRLPMIQLSQ 1398
Qy 1953 YVYLHQCVDRDLARKLRSEQENLFFIYE-----NVN----- 1985
Db 1399 YIFLHSCLLNKLSPSDASDSGPI-PWMNPAQACAKAANANAGFLKEYELLKQAIKDE 1457
Qy 1986 -----PEYHRDPVYSRH 1997
Db 1458 TGSLLSPDPVQNQSIASCH 1476

RESULT 4
US-11-143-984A-38
; Sequence 38, Application US/11143984A
; Publication No. US20060014180A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL HUMAN PHOSPHATASES
; FILE REFERENCE: D0072 DIV1
; CURRENT APPLICATION NUMBER: US/11/143,984A
; PRIOR FILING DATE: 2005-06-02
; PRIOR APPLICATION NUMBER: US 60/256,868
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: US 60/280,186
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/287,735
; PRIOR FILING DATE: 2001-05-01
; PRIOR APPLICATION NUMBER: US 60/295,848
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/300,465
; PRIOR FILING DATE: 2001-06-25
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 38
; LENGTH: 1711
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-11-143-984A-38
```

```
Query Match 11.4%; Score 1198; DB 7; Length 1711;
Best Local Similarity 25.3%; Pred. No. 9.5e-12;
Matches 411; Conservative 256; Mismatches 657; Indels 302; Gaps 53;

Qy 477 HANETSLSIMQTPVAWEKXIISLADRDLLLLHKSLSKDAKE-----FTFTDLVP 527
Db 46 HGKSTSLFLSVAELGGFDYALSRSVD-----SSGSEGOQLQANTNSGFEFHGLVP 100
Qy 528 GRKYMATVTSIGSDLN-SSSVKGRVPAQVTDLHVANQMTSSLFNTWTOAQODVIFYQ 586
Db 101 GSRYQLKLVTRPCWQNVITITARTAPTVRGQLHSAQSPARLEASNSDAPGDQDSYQ 160
Qy 587 VLLIHH--ENVVIKNEGISSETSYSPHSLKSGSLYSVVVTVVSGGSISSRQVWVEGRTVPS 644
Db 161 LLLVHLESQTLACNVSVSPDTLSVSGDLLPGTYQVLEVITWAGSLHAKTSILQW-TEPV 219
Qy 645 SVSGVTVNNGRNDYLSVSHLVAPGDVDNVEVTLSDHKVQVQSLVIKSVRECSFSSLTP 704
Db 220 PDHLALRALGTS-LQAFWNSSEGATSFHMLTDLGGTNTTAVIRQGVSTHTFLHSP 278
Qy 705 GRLYTVTITTRSGKYENHSFQERTVPDKVQGVSVNSARSQY-----LRVSW 752
Db 279 GTPHELKICASAG-----PHQIWGFSATEWTVPSYPSDVLVTLPLNELWASW 325
Qy 753 VHATGDFDHYEVTIKNNKNIOTKSIKSENECFVQLVFGRLYSVTVTTKSGOYEANEQ 812
Db 326 KAGLGARDGY--VLKSLGPMESTSTLGPBECNAVFGPLPPGHVTLQLKVLAGPYDAWVE 383
Qy 813 GN-----GRTPPVKDLTURNSTEDLHVW--SGANGVDVQYEIQLLNDMKVFPFPHL 866
Db 384 GSTLAESAALPREVPGARL-----WLDGLEASKQPGRRALLYSD--DAPGSL 429
Qy 867 VNT-----ATEYFTSLTPCGQVKILVLTISGDVQOSAFIEGTVPSAVKNIHSPNGAT 921
Db 430 GNLVSPSGATHVPCGLVPGAHYRVDIASSTGDISQS--ISGTSPLPPOSLEVISRSSP 487
Qy 922 DSLTVNMTPGGVDVSYTSAFRH-SQKV--DSQTIPIKHVFETHRLEAGEYQIMIAS 978
Db 488 SDLTIANGPAPQLEGYKVTWHQDGSQSPGLVDLGPDTLSLTLSKSLVPGSSVTSAWA 547
Qy 979 VSGSL-KNQINNVGRVTPASVQGVADNAYSSVLSVSWQKAAGVARYDILLTENGIL 1037
Db 548 WAGNLGSDSQKIHSCSTRPAPPTNLSLGFAGHQAALKASWYHPGGRDAFHLRLYRLPLT 607
Qy 1038 LRNTSE-PATTQKHKEFDLTPGKKYKIQLTVSGGLFSKBAQTEGRTVPAAVTDLRTEN 1096
Db 608 LESEKVLPREAQNFQVQAGLTAQCFQVQJLTLGWSERSSANATGWTTPSPAPTLVNTSD 667
Qy 1097 STRHLSFRWTASEGELSWYNIFLYNPDGNLQERAQVDPLVQSFSPQNLLGRMYKMVIT 1156
Db 668 APTQLQVMAHVPGRSRYQVTLTYQ-ESTRTATSIMGPKEDGTSFLGLTPGTYKKEVIS 726
Qy 1157 HSELGNESIFGRTVTPASVSHLGRNNTTDSLWFWNSPASGDFPFYELLINPNCTKK 1216
Db 727 WAGPL-----YTAAANVS-----AWT-----YPLI-----PN----- 748
Qy 1217 ENWKDLDLTEMRFQGLVPGKRYVYLVVWTHSGDLSNKTVAESRTAPSPSPSLMSADIANTS 1276
Db 749 -----ELLVSMQAGS-----AVNLAWSGP-LGQGACHAQLS 780
Qy 1277 LA--ITWKGPDPDWTYNDYDFELWLPDRDALTVFNPNYNNRSEGRIVYVGLR---PGRSYQFN 1331
Db 781 DAGHLSWEQ-----LKLGOELFWLRLDTFGHTISMS 812
Qy 1332 VKTVSGDSWTKYSKPIFGS-----VRTKPKIQLNLCRPNQSTAIACSWTPPDSDFGYS 1386
Db 813 VRCRAG-----PLQASTHLVLSVEPGFVDLCHPE-ATYALNMTWMPAGDVDCVL 863
Qy 1387 IEC-RKMDTQEVFSRKLKEKSLNIMLVPHKRYLVSLKVQSGAGT-SEVVEDSTITM 1444
Db 864 VVVERLVPGGTHFVQVNTSGDALLPLNLMPTTSTRLSLTVLGRNRSWRASVLCVSTS 923
Qy 1445 IDRPPPP-----PPHVRNEKDVLSKSSINFTVNCWSFSDTNQAVKYFTVVV----- 1492
```





```
Db 749 ADPEVFCCQVGGSQKTKLQEPVAVSS-----HVVTI----- 780
Qy 1545 NIKLGAEMESLGGKRDPTQKFCGDKPLKPHPTAYRISIRAFQTFDLKFTKPLYSDF 1604
Db 781 -----SSLLPATAYNCSTVFSH-----DPSV-PTF 806
Qy 1605 FSL-PITTESPLFGAIGVS-AGLFLIGMLVAVVALLICROK-VSHGRERPSARL----- 1657
Db 807 IAVSTMVTEMNPNVWISVAILSTLLIGLL--LVTLILRKHQLQWARECGAGTFVNF 864
Qy 1658 SIRDRLPSVHLNLGQGNKRTSPKINQEGHFMKLOADSNVLLSKEVEELKOVGNQ 1717
Db 865 SLEDDGKLPNWMSKNGKLKTNVQLDDFDAYIKOMAKDSYKFSLQFEELKLGDI 924
Qy 1718 SCDIALLPENRGNKRYNNILPYDATRVKLSNVDDPCSDYINASYIPGNFRREVITQG 1777
Db 925 PHFAADLPNRCRKRYNLPYDFSRVNLVMSNEEGADYINANYIPGYNPQBIATQG 984
Qy 1778 PLPGTKDDFWKMWQEQNVHNIWVMTQCVKGRVKCDHYWPAQDSLYYGDILILQMLSESV 1837
Db 985 PLPETRNDFWKWLQKQSQIIVMLTQCNEKRRVKCDHYWPFTEPIAYGDITVEMISEEE 1044
Qy 1838 LPEWTIEFKICGEQDLDAHLRIHFFHYTWPDHGV-ETTSQSLIQVTRVRYINRSP 1895
Db 1045 QDDWACHRFRI---NYADEMQDVNMFNYTAMPDGHGVPTANAESILOFVHMVRQQAATKS- 1100
Qy 1896 GAGPTVVCHSAGVGTGTFTIALDRILQOLDSKSDVIYGAVHDLRLHRVHMVQTECOVY 1955
Db 1101 -KGMITHCSAGVGTGTFTIALDRLLQHIRDHEFVDILGLVSEMRSTMSVQTEQYIF 1159
Qy 1956 LHQCRVDVLRARKLRSEQENPLFPPIYENVN 1985
Db 1160 IHQCVQLMMKKK---QQFCISDVIYENVN 1186
```

## RESULT 6

```
US-11-000-463-250
; Sequence 250, Application US/110000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785C1P4CN
; CURRENT APPLICATION NUMBER: US/11/000,463
; CURRENT FILING DATE: 2004-11-29
; PRIOR FILING DATE: 10/291,265
; PRIOR FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/922,279
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 250
; LENGTH: 1907
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-000-463-250

Query Match 7.9%; Score 832; DB 7; Length 1907;
Best Local Similarity 23.5%; Pred. No. 4.le-47;
Matches 408; Conservative 244; Mismatches 673; Indels 410; Gaps 75;

Qy 521 TTFDTLPGKRYKMATVTSISGDLKNSSSVKGRTPVPAQVTDLHVANQGTSSIFTNWTCAQG 580
Db 10 TMVPLVPALVMGLVAGAHGDSK-PVFIK---VPEDQT-----GUSGGVASVFCQATG 58
Qy 581 D-----VEF-----YQVLLIHENVVKNESISSETSRYSFHSLS 615
Db 59 EKPRIITMMKKGKVSQRFEVIEDGAGSVLRIOPLRVORDEAIECTATNSLGEINT 118
Qy 616 GSIYSVV-VTTVSGGTSRQVWVEGRTPVPSVSGVTNNSGRNDYLSVSMVAPGDVDNY 674
Db 119 SAKLSVLEEBQLPPGFPSIDMGFQLKVEKARTATMLCAAGGNPDPEISWFKOFLPYD-- 176
Qy 675 EVTLSDHGKVQ---SLVIKSVR-----ECSSFSS-----LTPGRLYVTITTSRGK 718
Db 177 --PATSGRIKQRLSGALQIESSESDQKCECVATNSAGTRYSAPANLY-VRVRVAPR 233
Qy 719 YENHSFQERTVPDKVQGVSVNSARSVDYLRVSWHATGDFHYEVTIKN----NNFIQ 774
Db 234 FSIPSSQEVMPGGSVNLTCVAVGAPMPY--VKMMGA-----EELTKEDEMPVGRNVLE 286
Qy 775 TKSIPKSENCVFLVPGRLYSVTTTKSGQVANEQNGRTIPEPVKDLTLNRSTED 834
Db 287 LSNVRSAN-----YTCVAISSLGMIETAQVTVKALPKPDLVVTETTATS 334
Qy 835 LHVTWSGANGD-VDOYEIOLLENDMKVFPFHLVN--TATEYRFTSLTPGROYKILVTI 891
Db 335 VTLTWDSGNSEPTYGIQ--YRAAGTEGPFQEVGVATTRYISIGLSPSEVAFRVIAV 392
Qy 892 S--GDVQQSAFIE---GFTVPSAVKNIHISPNGATDSLTVNTPGGGDVDSYTVSAFRHS 946
Db 393 NSIGRGPPSEAVRARTGEQAPSPRRVQARMLSASTMLVQWEP--PEEPNGLVRGVYV 450
Qy 947 QKVDSTQTIKHFVEH-----TFHLEAGEQYQIMIASVG-----SLKNQIN 988
Db 451 YTPDSRRPNNAWHKHNNTDAGLLTTVGSLLPFGITYSLRVLAFTAVTAVGDPSPFTIQVKTQ 510
Qy 989 VVGRTPVASVOGVIAADNAVSSYSLVSWQKAAQVAERYDILLTENGILL----- 1038
Db 511 VPAQ--PADFOAEVE---SDTRIQLSW-----LLPQERIIMVELVYWAED 552
Qy 1039 -----RNTSEPAATTKQHKPEDLTPGKYYKIILTVSG---GLFSK--EAQTEGRVPAV 1088
Db 553 EDQGHKVTFDP--TSSVTLLEDLPDLYRFQLAARSDMGVGVFTPTTEARTAQSTPSAPP 610
Qy 1089 TDLRITENSTRHLSFRWT-----ASGELSWMYIFLYNPDGNLQERAOVDPLVQSFSQN 1143
Db 611 QKVMCSVMGSTTVRVSWVPPPADSRNGVITQYSVAVEAVDGEDGRHVDVDSIGREHSSWD 670
Qy 1144 LL---QGRMYKMWIVTHS--GELSNEFSIFGR-TPASVSHLGRSNRNTTDSLWPNW- 1194
Db 671 LVGLEKWTYRVWVRAHTDVGPDPESPVLVRDTEVDVPGPPRKRVEVEPLNSTAVHYWK 730
Qy 1195 ----SPASGDFDFYELILYNPNGTCKENMKDKL-----TEWR-----FQ 1230
Db 731 LPVPKQHGQIRGYQ-VTY---VRLNGEPGRGLPIIQDVMLAQAQWRPESEDIYETTIS 785
Qy 1231 GLVPGRKYVLWV-THSGDLSN---KVTAESTAPSPSLMSFADIANSTLAIWKGP- 1284
Db 786 GLTPETTYSVTAAATTKGDGARSKPKIVTTTGAVFGPRTMTTMTTAMNTAL-LQMHPPK 844
Qy 1285 -----PDWTDV----- 1290
Db 845 ELPGELLYRLQYCRADAEARPNTIDFGKDDQHTVTYGLHKGTYYIFRLAAKNRAGLGEFP 904
Qy 1291 -----NDFELQWLPR-----DALTVFNPNYNNRKSE 1315
```

Db 905 EKEIRTPEDLPSPFONLHVTLTSTTTELANDPVLAEGRNRIISYTVFRDINSQEL 964  
Qy 1316 GRI-----VYGLRPGRSYQFNKVTSGDSWKTY-SKPIFGSV--RTKPKD---IQNLH 1362  
Db 965 QNITTDTRFTLTGLKPDYTDIKVR-----AWTSKSGSGLSPSIQSRTMPVEQVFAKNFR 1019  
Qy 1363 CRPQNSTAICSWIPDSD-----PDGYSIECRKMDTQEVFEFSKLEKEKSLNIM 1413  
Db 1020 VAAAMKTSVLSWEPVDSYKSAVPPKILYNGQSVB---VDGHSM---RKL-----IA 1065  
Qy 1414 MLVPHKRY--LVSIKQVAGMAGTSEVEDSTITMDIRPPPPHVRVNEKDVLSKSSINF 1471  
Db 1066 DLQPNTEYSFVLMNREGSAGGLQHLVSIRTPADLLPHKPLPASAYIEDGRDLSNPHVQD 1125  
Qy 1472 TVNCSWPSDGTGAVKFTYVVREADGSDDELKPEQOHP-----LPSYLE-----YRH 1517  
Db 1126 PSLVRWF-----YIVVVPIDRVGGSMLTRPWSPTBELELLELAIEQGBEQRRRR 1177  
Qy 1518 NASIRVYQTYNFASKCAENPNSNSKFNILKGAEMESLGGKRDPTQOKFCGDLKPHYAT 1577  
Db 1178 ROAERL--KPYVAAQLDVL-----ETFTLGDKN--YRGFYNRPLSPDLSY 1220  
Qy 1578 RISIRAFQTFDEDLKEFTKPLYSDFP--SLPITTESEPLFGATEGVSAGLFLIGMLVA 1635  
Db 1221 QCFVLASLK-EPMDQKRVASSPYDEIVVQTPAQOQEEPEMLWVGVLAVILI--ILI 1277  
Qy 1636 VVALLICRQKVSHGRRPSAR--LSIRDRDPLSVH-----INLGQGNRRTSCPIK 1684  
Db 1278 VIAILLFKRKRTHS---PSSKDEQSIGLKSLLAHSSDPVEMRRLNYQTPCGRDHP-PIP 1333  
Qy 1685 INQFEGHPMKLOADSNYLLSKEYBELKDVGRNQSCDIALLENRGJNRNNILPYDATRV 1744  
Db 1334 ITDLADNIERLAKNDGLFSQYESI-DPGQOFTWENSLEVNPKPNRYANVIAYDHSRV 1392  
Qy 1745 KLSNVDDPCSDYNASVIGPNRRREVIYVQGPLRGTKDPMKMWQNVHNVMTQC 1804  
Db 1393 ILTSDGVGSDYINANTIDGRKQNIATQGPLFETMGDFWRWVMEQRTATVMMTRL 1452  
Qy 1805 VEGRVKCDHYWADQDLSYGLILQMLSESVLPETWIREFKI--CGEEOLDARLIRH 1862  
Db 1453 EEKSRVKCDQWPA-RGTETGLIQLVTLDTVELATYTVTRFALHKSSE--KRELQ 1508  
Qy 1863 PHYTWPHGVPEPTQSILQIVRVYRVYINRSP-GAGPTVHVCSAGVGRGTGFIALDRIL 1921  
Db 1509 PQFWAMPDHGVPEYPTPIALFLRRVKAC--NPLDAGPMVHVCSAGVGRGTGCFIVIDAML 1565  
Qy 1922 QOLDSKOSVDIYGAVHDLRLHVRHVMVOTECOVVYLHQ-----CVRDVLBARKL 1969  
Db 1566 ERMKHEKTVDIYGHVTCMRSORNYMVQTEDQYVFIHEALLAEATCGHTEVPARNL 1620

## RESULT 7

US-10-821-234-1635  
; Sequence 1635, Application US/10821234  
; Publication No. US20050255114A1  
; GENERAL INFORMATION:  
; APPLICANT: Labat, Ivan  
; APPLICANT: Stache-Crain, Birgit  
; APPLICANT: Andarmani, Susan  
; APPLICANT: Tang, Y. Tom  
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia  
; FILE OF INVENTION: 821A  
; CURRENT APPLICATION NUMBER: US/10/821,234  
; CURRENT FILING DATE: 2004-04-07  
; PRIOR APPLICATION NUMBER: US 60/462,047  
; PRIOR FILING DATE: 2003-04-07  
; NUMBER OF SEQ ID NOS: 1704  
; SOFTWARE: pt\_seq\_genes Version 1.0  
; SEQ ID NO 1635  
; LENGTH: 1897  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-10-821-234-1635

Query Match 7.9%; Score 826; DB 6; Length 1897;  
Best Local Similarity 23.5%; Pred. No. 1e-46;  
Matches 407; Conservative 244; Mismatches 670; Indels 410; Gaps 75;

Qy 525 LVPGRKYMATVTSISGDLKNSSSVKGRVTPAQVTDLHVANOGMTSSSLFTNNWTAQGD--- 581  
Db 4 LVPALVGLVAGAHGSK-PVFIK---VPEDQF-----GLSGGVASFVCQATGEKPK 52  
Qy 582 -----VEF-----YQVLLIHENNVWIKNESISSETSRYSFHSLSKSGSLY 619  
Db 53 RITWMMKKGKVSQRFEVIBFDDGAGSVLRIOPLRVQORDEALYECTATNSLGEINTSAKL 112  
Qy 620 SVV-VTTVSGGISRQVQVVEGRTPSSVSGVTVNNSGRANDYLSVSWLVAPODVNYEVL 678  
Db 113 SVLEEOQLPGFPIDMGPOLKVVEKARTATMLCAAGGNPDPEISWFKDFLPVD-----PA 168  
Qy 679 SHDCKVVQ-----SLVIAKSVR-----ECSPSS-----LTPGRLYTVTITTRSGKYENH 722  
Db 169 TSNRIKQLRSGALQIESSESDQGYECVATNSAGTRYSAFANLY-VRRVRVAPRFSIP 227  
Qy 723 SFSQERTVPKVOQVSVNSARSYDLRVSWVHATGDFDHYEVTIKK-----NNFIQTKSI 778  
Db 228 PSSQEVMPGGSVNLTCVAGAPMPY--VKWNGA-----BELTKEDEMPVGRNVLELSNV 280  
Qy 779 PKSNCECVFQVLVPGRLYSVTVTTKSGOYEANEQNGRTIPEPVKDLTLNRSTEDLHVT 838  
Db 281 VRAN-----YTCVAISSLGMIEATAQVTVKALPKPIDLVVTTTATSVTLT 328  
Qy 839 WSGANGD-VQOYEIQLLFNDMKVPPPHLYN--TATEYRFTSLTPGRQYKILVLTIS--G 893  
Db 329 WDSGNSEFVTVYGIQ--YRAAGTEGPFQEDVGAVTTRYSGISGLSPFSEYAFRVLAVNSIG 386  
Qy 894 DVQOQSAFIE--GFTVPKSAVKNIHISPNGATDSLTVNMTPGGQVDVSVTSFAFHSOKVD 950  
Db 387 RGPSEAVRARTGEQAPSSPRRQVQARMLSTMLVQWEP--PREPNGLRGYRVYTPD 444  
Qy 951 SQTIPKHVFEH-----TFHRLAGEQYQIMTASVSG-----SLKQINVVGR 992  
Db 445 SRRPNAWHKHTDAGLLTTVGLPGITTSRLVLAFTAVGDGPPSPPTIQVKTQGVPAQ 504  
Qy 993 TVPASVQGVADNAYSYSLSIVSWQKAQVARYDILLITENGILL----- 1038  
Db 505 --PADFOAEV-----SPTRIQLSW-----LLPQOERIIMVELVYMAAEDDQ 546  
Qy 1039 -RNTSEATTQKHKFEDLTGKTKYKIQLTVSG--GLFSK--EAQTEGRVTPAAVTDLR 1092  
Db 547 HKVTFDLP--TSSYTLDELKPDTLRFQLAARSDMGVGFPTIERTAQSTPSAPPQVM 604  
Qy 1093 ITENSTRHLSFRWT-----ASEGELSWYNIFLYNPDGNLQBARQVDPVLQVFSFQNL-- 1145  
Db 605 CVSMGSTTVRVSVVPPPADSRNGVITQYSVAHEAVDGEDRGHRHVVDGISREHSDVLVL 664  
Qy 1146 -QGRMYKNVIVTHS--GELSNESIFGRT---VPASVSHLGRSNRTTDSLWFWN----- 1194  
Db 665 EKMTEYRVVWRAHTDVGPPESSPVLVRTDEDVSPGPRKVEVEPLNSTAVHVYMKLPVP 724  
Qy 1195 SPASGDFDFVELILYNPNGTKNWKDKDL-----TEWR-----FOGLVP 1234  
Db 725 SKQHGQIRGQ-VVY-----VRLNGEPRGLPIIQDVMLEAQWPRSESEDYETISGLTP 779  
Qy 1235 GRKYVLWV--THSGDLSN---KVTAESRTAPSPPLSMFADIANSTLAITWKP----- 1284  
Db 780 ETTYVTVAAATTKGDGARSKPKIVTTTGAVGPRPTMMISTMTANTAL-LQWHPKELPG 838  
Qy 1285 -----PDWTDY----- 1290  
Db 839 ELLGYRLQYCRADARPNTIDFGKDDQHFVTGLHKGKTYIFRLAAKNRAGLGEFEKEI 898  
Qy 1291 -----NDFELQWLP-----DALTVNPNYNNRKSSEGI- 1318  
Db 899 RTPEDLPSPFQNLHVTLTSTTELADPPVLAEGRNRIISYTVFRDINSQELQNT 958



Db 121 KEIDGIATTR-YSVAGLSPYSDYEFRRVAVANNIGRPASEPVLQTSEQAPSSAPRDVQA 179  
Qy 1094 TENSRLHSLFWTASE---GELSWYNIFLYNPD-----GNIOERAQVDPLVQSPFSQNL 1145  
Db 180 RMLSSSTILVQWKEPPEPQIQGYRVY-YTMDPTQHVNNMMKHVADSDQITIG--NLV 236  
Qy 1146 QGRMYKVVIVTHS---GELSNESFIFGRT-VPASVSHLRGNSNTTDSLWFNWSPASGD 1200  
Db 237 POKTYSVKVLAFTSIGDPLSSDIQVITQTVGPQPLNFKAPESET-SILLSTWPPRSD 295  
Qy 1201 -FDFEYLILNPNGTCKENWKDKOLTWRFOGLVGRKYVLWVTHSGDLSNKTAE--S 1257  
Db 296 TIASVELVYRDGDQGEORITIEPCTSYRLQGLXPNLSLYPRLSARSPQGLGASTAISA 355  
Qy 1258 RTAPSPSLMSFAD-----IANTSLAITWKGPDPMDTYNDYDFELQWLPDRDALTVFNPN 1311  
Db 356 RTMQS-----MFAKNFHVKAVMKTSVLLSWEPENYSAMPFKILY---DQKMWEEVDG 407  
Qy 1312 RKSEGRIVYGLRGRSYQFNKTVTSGDS-----WKTYSKPIFGSVRTKPKIQNLHCRPQ 1366  
Db 408 RATO-KLIVNLKPKESYSF-VLTNRGNSAGLQHRVTAKTAPDLVLRTPKAPFIGNL--D 463  
Qy 1367 NSTAIACSWIPDDDFGYSIECRKMDTQEVFESKLEKESLNNMMLVPHKRYLVSIK 1426  
Db 464 GMITVQPLDVPANENIKGYI-----IIVPLKK-----491  
Qy 1427 VQSGMTSEVVEDSTITMIDRPPPPHPIRVNE---KDVLSKSSINFTVNCWFSFDTNGA 1484  
Db 492 --SRG-----KFIKPWSPDEMDLDELKELSRKRSIRYG-----525  
Qy 1485 VKYTVVVRADGDELKPEQOHPPLPSYLEYRHNASIRVQNTYFASKCAENPNSNSKSF 1544  
Db 526 -----REV-----ELKP-----YIAAFDVLPT-----543  
Qy 1545 NIKLGAEMESLGGKRDPTQOKFCGPKLPHPTAYRISIRAFQTLDFEDLKEFTKPLYSDTF 1604  
Db 544 EFTLGDD-KHYGG-----FTNKQJOSGOEYVFFVLAV--MDHAESKMATSPYSDPV 592  
Qy 1605 FSL-----PIITSESEPLFGAIEGVSAGLFLIGMLVAVVALLICRQVSHGRERPSARLSI 1659  
Db 593 VSMDLDPQITDDEGLIWWVGPVLAVFII---CIVIAILLYKKRAESERKS---SL 646  
Qy 1660 RDRPLSVH-----LNLGQGNKRTSCPIKINQFEGHFMKQADSNYLLSKYEYEE 1709  
Db 647 PMSKEVPSHPTDPVELRLNF-QTPGMASHPPPILELADHIERLKANDNLKFSQYES 705  
Qy 1710 LKDVGRNOSCDIALLPENRKNRYNNILPYDATRVKLSNVDDDDPCSDYINASYIPGNFR 1769  
Db 706 I-DPQOQFTWHSNLEVNKPKNRVANVIADHRSVLLSAIEGIPGSDYVNAVYIDGYRKQ 764  
Qy 1770 REYIVTQGPLPGTKDDFWKMWQVQVNHVNVVTCVKEGRVKCDHYWPAQDSLYYGDLI 1829  
Db 765 NAYIATQGSLETPETGDFWMLWEORSATVVMWTKLEERSRVKCDQYWPS-RGTETHGLVQ 823  
Qy 1830 LQMLSESVLPWMTIREPKICEEQLDHLRIHPHYTVPDGHVPETQSLIQFVRTVD 1889  
Db 824 VTLLDVTVELATYCVRTALY-KNGSSEKREVRQFTAPDPDGHVEHPPTPLAFLRRVKT 882  
Qy 1890 YINRSPGPTVHCHSAGVGTGTFFIALDRILQOLDKSDVDIYGAVHDLRLHVRHVQOT 1949  
Db 883 C--NPPDAGPMVHCHSAGVGTGTFCFVIDAMLERIKHEKTVDIYGHVTLMAQRNVQOT 940  
Qy 1950 ECQVYVLHQCVRDVL 1964  
Db 941 EDQVIFHDALLEAV 955

RESULT 10

US-10-821-234-1102

; Sequence 1102, Application US/10821234

; Publication No. US200502551141

; GENERAL INFORMATION:

APPLICANT: Labat, Ivan  
APPLICANT: Stache-Crain, Birgit  
APPLICANT: Andarmani, Susan  
APPLICANT: Tang, Y. Tom  
TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia  
FILE REFERENCE: 821A  
CURRENT APPLICATION NUMBER: US/10/821,234  
CURRENT FILING DATE: 2004-04-07  
PRIOR APPLICATION NUMBER: US 60/462,047  
PRIOR FILING DATE: 2003-04-07  
NUMBER OF SEQ ID NOS: 1704  
SOFTWARE: pt\_seq\_genes Version 1.0  
SEQ ID NO 1102  
LENGTH: 1452  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-821-234-1102

Query Match 6.9%; Score 718.5; DB 6; Length 1452;  
Best Local Similarity 22.4%; Pred. No. 1.1e-39; Indels 429; Gaps 42;  
Matches 269; Conservative 164; Mismatches 338

Qy 986 QINVVGRTVPAS---VQGVADNA-----YSSYSLIVSWQKAGVAERYDIL 1029  
Db 205 QCSAIGTAVAGDRLWLQIDVRDAPLKEIKVTSRRFIASFNVTTKRDAG---KYRCM 261  
Qy 1030 LLENGILLRWTS-----1043  
Db 262 IRTEGGVIGSYAELVVKPEVPPIAPQLASVGATYLIQNLNANSINGDGPVAREVEYC 321  
Qy 1044 -----PATYKQHKFEDLTPGKXKIQILTV---SGCLFSKEAQTEGRTVPA--- 1086  
Db 322 TASGWMNDROPVSTSYKIGHLDPDTEYBISVLLTRPEGGTGSPGALRTRTKCADPMR 381  
Qy 1087 AVTDLRITENSTRHLSFRMTASEGELSWYNI-----FLYNDGNLQERAQVDPL 1135  
Db 382 GPRKLEVEVKSQITIRW-----EPGYNVTRCHSYNLTVHYCYQVGQEQVREEV--- 433  
Qy 1136 VQSFSPONLQGRMYKVIIVTHSGELSNESFIFGRTVPASVSHLRGNSNRNTDLSLWFWNS 1195  
Db 434 --SWDTEN---SHPOHTITNLSPYTNVS-----456  
Qy 1196 PASGDDPFYELILYNPNGTCKENWKDKOLTWRFOGLVGRKYVLWVTHSGDLSNKTVA 1255  
Db 457 -----VKLILNPNPEGRKES-----QEL-----IVQDDEDLPQAVPT 487  
Qy 1256 ESRTAPSPSLMSFADIANSLAITWKGPDPMDTYNDYDFELQWLPDRDALTVFNPNNRKSE 1315  
Db 488 ESQGST-----FEEKIFLOWREPTQ--TYGVITLYEITYKAVSSDPDEIDLSNQ 535  
Qy 1316 -GRI-----VYGLRPGRSYQFNVK--TVSGDSWKYTSKPIFGSVRTKPKIQNLH 1362  
Db 536 SGRVSKLGNETHFLFFGLYPGTYSFTIRASTAGK-----FGPPATN-----577  
Qy 1363 CRQNSTAIACSWIPDSDFDGYSIECRKMDTQEVFESKLEKESLNNMMLVPHKRYL 1422  
Db 578 ---QFTTKISAPSM-----AYELETPNLQT-----600  
Qy 1423 VSIKVSAGMTSEVEDSTITMIDRPPPPHPIRVNEKDVLSKSSINFTVNCWFSFDTN 1482  
Db 601 -----DNTVTMLKP-----AHSR 614  
Qy 1483 GA-VKYFTVVVVRADGDELKPEQO-----HPLPSYLEYRHNASIRVQNTYFASK 1532  
Db 615 GAPVSVVQIVV-----EERPRRTKTKTEILKCYVPVPIHFQ---NASL-LNSQYFEA-- 662  
Qy 1533 CAENPNSNSKSFNIKLGAEMESLGGKRDPTQOKFCGPKLPHPTAYRISIRAFQTLFEDL 1592  
Db 663 -AEFPADSLQA-----AQPFTIGDNK--TYNGWNTPLLPYKSYRIYFOAASRANGETK 713  
Qy 1593 KETKPLYSDTFPSLPIITTESEPLFGAIE--GVSAGLFLIGMLVAVVALLICRQVSHGR 1650  
Db 714 IDCQVATKGAATPKPBPBEKQTDHTVTKIAGVIAGILLFVILFGLVGLVWNGKRLA--K 771

Qy	1651	ERPSARLSIRDRPLSLVH- :    : : : :: ::	-----LNLGQGNKRTSCP-----	1682
Db	772	KRKETMSSTRQEMTVMNSDKSYAEQGTNCDEAFSMDTHNLNGRSVSSPSSFTMKTTNT	831 :    : : : :: ::	
Qy	1683	-----TKINOFEGHFVK	1694	
Db	832	LSTSVPNYSYPDETHTWASDTSSLVQSHTYKKREPADVPTQGQLHPAIRVADLLQHITQ	891 :    : : : :: ::	
Qy	1695	LOADSNSYLLSKSEYEELKDVRNOSCDIALPENRGNKNRYNNILPYDATRVKLSNVDDDPCC	1754	
Db	892	MKCAEYGFGKEEYESFPE-GCSAPWDSAKDENRMKNRYNIIAYDHSRVRLQTIEGDTN	950 :    : : : :: ::	
Qy	1755	SDYNASYI PGNFRREYIVTQGPLKTOKDFWKVMVEQNHNIMVTQCVEKGRVKCDH	1814 :    : : : :: ::	
Db	951	SDYINGNIIDGYHRPNHYIATQCFMQETIYDFRWVWHENTASIIWMNLVEVGRVKCKC	1010 :    : : : :: ::	
Qy	1815	YWPADOBSLYGGDLILQMLSESVLPEWTIREFKICBEEQLDAH--RLIRHFHYTVPWDHG	1872 :    : : : :: ::	
Db	1011	YWPDDTE--IYKDIKVTLISTELLAEEVIIRTPAV---EKRGVHEIREIRQFHFTGWPDHG	1065 :    : : : :: ::	
Qy	1873	VPETTQSLIOFVRTVDYINRSGAGTYVHCAGVGRTGTFIALDRILLQOLDSKSVDI	1932 :    : : : :: ::	
Db	1066	VPHYATGLLGFLVRQVKS--KSPPSAGPLVVHCAGAGRTGCTFVIDIMLDMAEREQVVDI	1123 :    : : : :: ::	
Qy	1933	YGAHVHLRLHRVMVOTECOVYLHQCRDV-----LEARKLRSEQENLPFIYENVN	1985 :    : : : :: ::	
Db	1124	YNCVRELRRRVNMVQTEQGVFIHDAILAELCLCGDTSVPASQVRS-----LYYDMN	1175 :    : : : :: ::	
 RESULT 11				
US-11-080-991-22				
; Sequence 22, Application US/11080991				
; Publication No. US20050266437A1				
; GENERAL INFORMATION:				
; APPLICANT: Veiboy, Petter Ole				
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR				
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST				
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST				
; FILE REFERENCE: MRI-039				
; CURRENT APPLICATION NUMBER: US/11/080,991				
; CURRENT FILING DATE: 2005-03-11				
; PRIOR APPLICATION NUMBER: US/10/176,847				
; PRIOR FILING DATE: 2002-06-21				
; NUMBER OF SEQ ID NOS: 112				
; SOFTWARE: FastSeq for Windows Version 4.0				
; SEQ ID NO 22				
; LENGTH: 1463				
; TYPE: PRt				
; ORGANISM: Homo sapiens				
US-11-080-991-22				

```

Db      557  YPGTTSFTIKASTAKGFGPPVTRTIATKISAPMPE-----YDTDPLNETDITIVML 611
Qy      1476  SWFSDTNGAVKYFTVVVRADGSDLEKPEQQ-----HPLPSYLEYRINASIRVVQTNV 1528
Db      612  KPAQRCAPVSVQLVKE-----ERLQKRRADIIIECESVP--VSYRNASSLD--SLHY 663
Qy      1529  FASKCAENPNNSKSFNIKLGAEMESLGKRDPTQOKFCDGPLKPHATYRISRAPTOJLF 1588
Db      664  FAABL-----KPAANLPV-TQPFVTGDNK--TYNGYWNPLPLSKYSIVFQALSKAN 712
Qy      1589  DED-----LKEFTK-PLYS-----DTFSPILPITSESEPLGAE-----GVSG 1626
Db      713  GETKINCVRLATKAPGSAQVTFGTPLCLLTTCGASTQNSNTVEPEKQVNDTVKMAGVIAG 772
Qy      1627  LFLIGMLVAVVALLICROKVSH-----GRERPSAR 1656
Db      773  LLFPIILLGVMLTIKRRNAYSYYLSORKLAKKQKETSQAQREMGPVASADKPTTK 832
Qy      1657  LSTRD-----RPLSVHNLGOKNKR-----KTSCKPIKI-----NQPE---- 1689
Db      833  LSASRNDEGFSSSQSDVNGFTDGSRGELSQPTLTITQTHPYRTCDPVEMSYPRDQFLAIR 892
Qy      1690  -----GHEMKLOADSNYLLSKVEEELKDVGRNOSCDIALLPENRGKNRYNNILPYDATRV 1744
Db      893  VADLLOHITQMKGGQGYGFKEYEALPE--GQTASWDTAKEDNRNKNRYGNIISYDHSKV 951
Qy      1745  KLSNVDDPCSDYINASYIPGNFNREYIVTQGPLFGTKDDFWKMWVEQNVHNIVMVTCQ 1804
Db      952  RLLVLGDGPHSDYINANYIDGYHRPHRYIATQGPQMETVKDFPWRMIWQENSASIVMVNVL 1011
Qy      1805  VEGGRVKCDHYNPADOOSIYIGDLLLOMLSESVLPBWTITREFKICBEOQLDAH--RLIRH 1862
Db      1012  VEVGRVKCRYVDPDDTE--VYGDIKVTLTETPLAEYVIRTFV---QKKGYHEIRLERL 1066
Qy      1863  FHYTVPDGHVPTTQSLIQFVRTVDYINRSPGAGPTVVHCSAGVGTGTGTGTGFIALDRILQ 1922
Db      1067  FHTISWPDHGVCYATGLGFRQVK-FLN-PPEAGPIVHCSAGAGRTGCGFIADTWLD 1124
Qy      1923  QLDSKDSVDIYGAVHDLRLHRVHMVQTECQYVYLHQCV 1960
Db      1125  MAENEGVWDIFNCVRELRAQRVNLVQTEQQYVFWHDAI 1162

RESULT 12
US-11-097-728-2
; Sequence 2, Application US/11097728
; Publication No. US20050260132A1
; GENERAL INFORMATION:
; APPLICANT: Erik Foehr
; APPLICANT: Sabine Muller
; APPLICANT: Daniel J. Chin
; APPLICANT: Mirella Gonzalez-Zulueta
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES DIRECTED TO RECEPTOR PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASE ZETA
; FILE REFERENCE: AGYT-006CIP
; CURRENT APPLICATION NUMBER: US/11/097,728
; CURRENT FILING DATE: 2005-03-31
; PRIOR APPLICATION NUMBER: 10/652,981
; PRIOR FILING DATE: 2003-08-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2314
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-097-728-2

```

Query Match 6.6%; Score 687.5; DB 7; Length 2314;  
Best Local Similarity 21.2%; Pred. No. 2.7e-37;  
Matches 466; Conservative 321; Mismatches 763; Indels 653; Gaps 106;  
Qy 176 GSKYNTAITAVSGGKRFSFVYNTGTVSPVKDIGISTKANSLLLSWHSHGSGNVRYRLM 235

Db 37 GWSYTGALNQWKK---YFCNSPKQSPINIDELTQVNV-----NLKULFKQ 83  
Qy 236 LMDKGL-----VH---GGVVDKHTS-YAFHGLSPGYLYNLVMTTEAAGLQRYRWKLVRTA 288  
Db 84 GWDKTSLENTPIHNTGKTVEINLTNDYRVSG-----GVSEMVFKASKITFWH----- 130  
Qy 289 PMEVSNLKVTNDGSLTSLKVKWQRP-----GNVDSYNTLSHKGITIKESRVLAP 338  
Db 131 -----GKNMSSDGGSEHLS--EGQFPLEMOYICFDADRPFSSFEBAVKGKGLRALSILFE 184  
Qy 339 WITETH--FKELVPGRLYQVTVSCVS--GELSAQMAVGRFPDKVANLEANN----- 387  
Db 185 VGTEENLDFKAIIDG-----VESVSFRGQAALD-----PILLNLLPNSTDKYIY 231  
Qy 388 NGRMRSVLVWSPAGDWEQVRIILLFNDSSVLLNITVCKBETQ-----YVMDTGLVLP 440  
Db 232 NGSULT-----SPCTDITVDN--IVFKD-----TVSISESQLAVCEVLTMQOSGVVM 276  
Qy 441 GRQYEVEVIVESGNLKNSERCQGRTPVLAVL-----QLRVKHANET 481  
Db 277 LMDY-----LONNFRQOYKFRQVSESYTGKEIHEAVCSSEPEVQADPENYT 326  
Qy 482 SLSIMWOTPVAEWKEYIISLA-----DRDILLHKSLODAKEF--TFDVLVPGRYMA 533  
Db 327 SLLVWERPRVYDTMIEKFAVLVQOLDGEDQTKHEFLTGYQDLGAILNLLFNMSYVL 386  
Qy 534 TVTISIGDLKSSSVKGRTPVPAQVTDLHVANQGM-----TSSLFTNMTQAGDVEFY 585  
Db 387 QIVAIK-----TNGLYGKYSQDLIVDMPDNPDLFPPELLIGTEIIEKEBEGK-DIE-- 438  
Qy 586 QVLLIH-----ENVVINKESISSETRYSHLSKSLYSVVVTVS-----GG 629  
Db 439 EGAIWNPGRDSATNQIRKQEPQISTTHYN-----RIGTKYNEAKTNRSPTRGSEFSGKGD 494  
Qy 630 I-----SSROVVEGRTV-----PSSVSGVTVN-NSG-----RNDYLSVS 663  
Db 495 VPNTSLNSTQPVTKLATEKDITSQVTELPHTVEGTSASLNGSKTVLRSPHNLS 554  
Qy 664 WLVAPOVDNVEVTLSDHDKVQVSLVIAKSVRECS-----FSSL 702  
Db 555 GTAB--SLNTVSIIEYEESLTSFKLDGTGAEDSSGSPATSAIPFISENISQVIFSE 612  
Qy 703 TPGLL-VTVTI--TTRSGKYENHFSQERTVPDKVQGVSVNSARSYDLRVSWVHATGDF 759  
Db 613 NPETITVDVLIPESARNASEDSTSGSESLKDPSEMGV-----WPPSSTDI 660  
Qy 760 DHYEVTIKNNFIQTK--SIPKSENECVFVLVPGRLYSYVTVTKSQYEAN--EQGN 814  
Db 661 TAQPDVSGGRESFLOQNTYETIRVDESE-----KTKS--FSAGVMSQG- 702  
Qy 815 GRTIPEPVKOLTLNRSTEDLHVTSWANGDQVQEIQLLFNDMKVPPPHLVNTATEYR 874  
Db 703 -----PSVTDLEMPHYST-----PAYEP-----TEVT 724  
Qy 875 FTSITPGROYKILVLT-----SGD--VQOASAFIEGP--TVPSAVKN--THIS 916  
Db 725 PHAFTPSSROQDLVSTVNVVYSQTPQVYNGETPLQPSYSEVEFPLTELLDNOILNT 784  
Qy 917 PNGATDSLTVNWTGGGDVD-----SYTVSAFRHSQKVDSTQIP 955  
Db 785 PAASSSDSALHATVPFVSVDVSFESILSSYDGAFLPLFPSSASFSEFLHHTV-SQILP 843  
Qy 956 KHVEFTPHRLEAGEQYQIMTASVSGLKNQINVVGRTPVPSVQGVLTADNAYSYSILVS 1015  
Db 844 Q-----VTSATEDKVPILHA--SLPVAGDGLLLEPSLAQYSDVLS 881  
Qy 1016 WQKAAGVAERYDILLTENGILLR-----NTSEPAITKQHKFEDLTPGKYKI----- 1063  
Db 882 TTHAASETLBEG-----SESGVLYKTMFMSQVEPPSSDAMMHARSSGPEPSYALSDNEGSQ 937  
Qy 1064 QILTVS-----GGLFSEKAEQTEGRTPAAVTDLRITENSTRHLSFRWTA 1107

Db 938 HIPTVSYSSALPVHDSVGVTVYQGSLSFGPSH-----IPIPKSSL-ITPTASLLQPTHALS 991  
Qy 1108 SEGELSWN-----IFLYNPDGNLQERAQVDPVQSFOQLNLOGRMYKMWLVTHSGELSN 1163  
Db 992 GDGEWSGASSDSEFLLPDDTDLTALNITSSPVSAEFTYTTTSVFGDDNK-----ALSK 1043  
Qy 1164 ESPTFGR-----TVPA-----SVSHLGRSNRNTTDSLWFNFWSPASGDF 1201  
Db 1044 SEIYIGNETELQIPSFNEMVYPSSEVPMYDNNVKNLNASLQETSIVI-----SSTKGMF 1099  
Qy 1202 DFYELIILYNPGTKKENWK--DKDLTEWRFFQGLVPGRYKYLWVVTH-----SGDLSNK-- 1252  
Db 1100 -----EGSLAHTTTTKVDFHEISO-----VPENNFSV-QPTHVTVSQASGDTSLKPV 1143  
Qy 1253 VTAESERTAPSPPS-----LMSFADITANTSLAITWKGPPDWTYDNDF--ELQWLP 1299  
Db 1144 LSNASEPASDPASSEMSPSTQLLFYETSASEFTEVLLOPFOAQSDVDTLLKTLVPAVP 1203  
Qy 1300 RDALTVPNP-----YNNRKSSEGRIVYGLRPGRSYQFNKVTYSGDSMKTYSK 1345  
Db 1204 SDPILVETPKVDKISSTMLHLIIVSNSASSENML-----HSTSVPDFVDSPTS-HMHS 1255  
Qy 1346 PIFG-SVRTKPDQIQLNLCRPNQSTAIACSWIPPDSDFDGYSIECRKMDTOEVE--FSRK 1402  
Db 1256 SLOGLTISYASEKVEPVLLKSESHQVVPVSLYSNDELFOPTANLEINQAHPPKGRHVATP 1315  
Qy 1403 LEKEKSLLNMM--LVPHRYKLVISIKVQSGAMT-----SEVVEDSTIIMIDRPPPPPHIR 1456  
Db 1316 VLSIDEPLNTLINKLIHSDEILTSTKSVTKVPAGIPTVASDFTFVSDHSPVINGHVA 1375  
Qy 1457 VN-----EKDVLISKSSINFVNCWFSFDTNGAVKYFTVVV-----READSGDELKPEQ 1505  
Db 1376 ITAVSPHRDGSVSTKLLFPKAT--SELSHSAKSDAGLVGGGEDGDTDDGDDDDDRDS 1433  
Qy 1506 Q-----HPLPSYLEVR-----HNASIRVYQNTYFASKCAENPNSKSNFKLKG 1549  
Db 1434 DGLSIHKCMSCSSVRESQEKVMDSDTHENSL-MQONNPISYSLSESEEDNRVTSVSD 1492  
Qy 1550 AE--MESLOGK--RDPTQOKPCDG-----PLKPHYATYIRISITRAFTQLDFED 1591  
Db 1493 SQTGMDRSPGKSPANGLSQKHNDGKEENDITQTSALLPLSPES-----KAWAVLTSDE 1546  
Qy 1592 -----LKE--FTKPLYSYT-----FFSLP--ITTSEP 1615  
Db 1547 ESGSGQTSLSNENETSTDFSFADTNEKDADGILAAAGDSEITPGFQSPSTSVTSENSE 1606  
Qy 1616 LF-----GAIEG-----VSAGLFLIGMLVAVVALLICROK- 1645  
Db 1607 VFHVSEABASNSHESRIGLAEGLESEKAVIPLVIVSALTFI--CLVVLVGLIYWRKC 1664  
Qy 1646 --VSHGRERPSARLSIRDRPLSVHLNLQKGNKRTSCPIKINQFEGHFMKLQADSNYLL 1703  
Db 1665 FQTAHFYLEDSTSPRVISTPTPIFPISDDVG-----AIPKHPKHVADLHASSGF-- 1716  
Qy 1704 SKEYBELKOVGRN-QSCDIAL-----LPENRGKRYNNILPYDATRVKLSNV--DDD 1752  
Db 1717 TEBFETLKBEFYQVQSCVTDLGITADSSNHPDNKKNRYINIYAYDHSRVKLAELAKDG 1776  
Qy 1753 PCSDYINASYIPGNPFREYIVTQGPLGTDKDPFKVMWQONVHNIYVMTQCVSKGRVKC 1812  
Db 1777 KLTDYINANYVDGYNRPKAVIAAQGLKSTAEDEFWRMIWEHNEVIYMTNLVBEGRKRC 1836  
Qy 1813 DHYWPAQDQSLYTGDLILQMLSESVLPETIREF-----KICGBEQDAHLRIRHFY 1865  
Db 1837 DQWPAID-GSEYGNFLVTKSQVQLAYYTVRNFTLNTKIKKGSQGRPSGRVVTVQVY 1895  
Qy 1866 TVMPDHGVPTTQSLIOFVTVRDYINRSPGAGPTVHCSAGVGRGTFTFALDRILQOLD 1925  
Db 1896 TQWPDMDGVPEYSLPVTFRKAA-YAKRH-AVGPPVWHCSAGVGRGTGYIVLDSMLQOIQ 1953  
Qy 1926 SKDSVDIYGAHDLRLHRHVHMQVTECOYVYLHQCVRDVLARK 1968  
Db 1954 HEGTVNIFGLKHIRSORNYLVQTEEQYVFIHDTLVEAILLSKE 1996

## RESULT 13

US-11-097-728-6

; Sequence 6, Application US/11097728

; Publication No. US20050260132A1

; GENERAL INFORMATION:

; APPLICANT: Erik Foehr

; APPLICANT: Sabine Muller

; APPLICANT: Daniel J. Chin

; APPLICANT: Mirella Gonzalez-Zulueta

; TITLE OF INVENTION: MONOCLONAL ANTIBODIES DIRECTED TO RECEPTOR PROTEIN

; FILE OF INVENTION: TYROSINE PHOSPHATASE ZETA

; FILE REFERENCE: AGYT-006CIP

; CURRENT APPLICATION NUMBER: US/11/097,728

; CURRENT FILING DATE: 2005-03-31

; PRIOR APPLICATION NUMBER: 10/652,981

; PRIOR FILING DATE: 2003-08-28

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 2353

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: VARSPLIC

; LOCATION: (1)...(2353)

; OTHER INFORMATION: FTP-zeta SM2 23a exon variant

US-11-097-728-6

Query Match 6.6%; Score 687.5; DB 7; Length 2353;

Best Local Similarity 21.2%; Pred. No. 2.8e-37;

Matches 466; Conservative 321; Mismatches 763; Indels 653; Gaps 106;

Qy 176 GSKYNIAITAVSGGRFSVYNGSTVPSPVKDIGISTKANSLLISWHSNGNVERVRLM 235  
Db 37 GWSYTGALNKNMGK---YPTCNSPKQSPINIDEDLTQVNV-----NLKLLKQF 83  
Qy 236 LMDKGLI---VH---GGVDVKHATS-YAFHGLSPGYLYNLTVTEAAGLQNYWKLVRTA 288  
Db 84 GWDKTSLENTFIHNTGKTVEINLTNDYRVSG-----GVSEMVFRASKITFWH----- 130  
Qy 289 PMEVSNLKVTNDGSLTSLKVKWQRP-----GNVDSYNITLISHKGTIKESRVLAP 338  
Db 131 ----GKCNMSSDGEHSL--EGQKFPLEMOIYCFDADRPFSSPEAEVKGKGLRALSLTPE 184  
Qy 339 WIETH--FKELVPGRLYQVTVSCVS--GELSQAQMAVGRTFPDKVANLEANN----- 387  
Db 185 VGTEENLDFAIIDG-----VESVSFRGQAALD-----PFILINLLPNSTDKYIY 231  
Qy 388 NGRMRLSVVSWSPGADWEQYRILLFNDVSVLLNITVGKEETO-----YVMDDTGLVP 440  
Db 232 NGSLT-----SPCTDTVDW--IVFKD-----TVSISEQLAVFCEVLTWQSGYVM 276  
Qy 441 GROVEVEVIVESGNLKNRSRCQRTVPLAVL-----QLRVKXANET 481  
Db 277 LMDY-----LQNNFREQQYKSRQVFSYVTGKEETHEAVCSSEPENVOADPENYT 326  
Qy 482 SLSIMQTPVAEKEYIISLA-----DRDLILHLSLSKDAKEP--TFTDLPVGRKYMA 533  
Db 327 SLLVTERPRVVDYTMIEKFAVLYQLDGDGDTKHEPLTDGYQDLGAILNLLPNMSYVL 386  
Qy 534 TVTISGDLKNSSVGRTPVPAQVTDLHVAQNM-----TSSLFTNTWTOAQGVFEY 585  
Db 387 QIVAIIC-----TNGLYKYSQDIIVDMPTDNPBLDLPPELIGTEETIKKEEGK-DIE-- 438  
Qy 586 QVLLIH-----ENNVIKNESISSETSRYSFHSLSGSLYSVVVTVTS-----GG 629  
Db 439 EGAIVNPGRSATNQIRKEPQLSTTHYN-----RIGTKYNEAKTNRSPTRGSEFSKGD 494  
Qy 630 I-----SSRQVVVEGRV-----PSSVSGVTVN-NG-----RNDYLSVS 663  
Db 495 VPNTSLNSTSQPVTKLATEKDISLTSQTVELPHTVEGTSASINDGSKTVLRSPPHNL 554

Qy 664 WLVPAGDVNDYEVTLSDHGKVKQSLVIAKSVRECS-----FSSL 702  
Db 555 GTAE--SLNTVSTITEYEESLLTSFKLDTCAEASSGSSPATSAIPFISENISQYIPSE 612  
Qy 703 TPGRLL-YTVTI--TTRSGKYENHFSQERTVPDKVQGVSVNSGARSDDLVSVMVHATGDF 759  
Db 613 NPETITVDVLIPESARNASEDSTSSGSESLKDPMEGNV-----WFPSTDI 660  
Qy 760 DHYEVTIKNNPIOTK--SIKSENECVFVLVPGRLYSVTVTTKSGQYEAN---BOGN 814  
Db 661 TAQPDVSGGRESFLOQNYTEIRVDESE-----KTKKS--FSAGFVMSQG- 702  
Qy 815 GRTIPEPVKDLTLNRNSTDLHVTWSGANGVDQYEIQLFNDMKVFPFPHLVNTATEYR 874  
Db 703 ----PSVTDLEPHYST-----FAYFP-----TEVT 724  
Qy 875 FTSLTGPRQYKILVLTIT-----SGD--VQOSAFIEGF--TVPSAVKN--IHIS 916  
Db 725 PHAFTPSSRQODLVSTVNVVYSQTTQPVYNGETPLQPSYSSEVFPLVTLPLLDNQILNTT 784  
Qy 917 PNCATDSLTVNWTGCGDYD-----SYTSAFRHSQKVDQOTIP 955  
Db 785 PAASSSDSALHATFPFSDVDVSFESILSSYDGAAPLFPSSASFSELFRLHRTV-SOILP 843  
Qy 956 KHVFEHTFHRLEAGEQYQIMIASVGSGLKNQINNVGRTVPASVQGVVIADNAYSYSILVS 1015  
Db 844 Q-----VTSATESDKVPLHA---SLPVGAGDILLPELSLAQYSDVLS 881  
Qy 1016 WOKAAGVAERYDILLTENGILLR-----NTSEPATTKQHKFEDLTPGKKYKI----- 1063  
Db 882 TTHAASETLEFG---SEGVLVKTLMFQVPEPPSSDAMHARSSEPESYALSDEGSO 937  
Qy 1064 QILTVS-----GGLFSKAQTEGRTVPAAVTDLRTENSTRHLSFRWTA 1107  
Db 938 HIFTVSYSSAIPVHDSVGVTYQGSFSPSH-----IPIPKSSL-ITPTASLLQPTHALS 991  
Qy 1108 SEGELSWYN---IFLYNPNGNLQRAQVDPLVQSPFQNLQGRMKVMKIVITHSGELSN 1163  
Db 992 GDGENSGASDSEFLPDPTDGLTALNISPVSVAEFTYTSVFGDNK-----ALSK 1043  
Qy 1164 ESFICR---TVPA-----SVSHLRGNNRNTTDSLWFWNWPASGDF 1201  
Db 1044 SEIYGNETELQIPSNEMVYSESTVMNMYDNVKNLASLQETSIVI-----SSTKGMF 1099  
Qy 1202 DFVELLYNPNGTKENWK--DKDLTEWRFQGLVGRKYVVLVWVTH-----SGDLNKK-- 1252  
Db 1100 -----PGSLAHTTTKVFDEHSQ-----VPENNFSV-QPHTTVSOASGDTSLKPV 1143  
Qy 1253 VTAESRTAPSPS-----LMSFADIANTSLAITWKGPDPDWDYDNF---ELQWLP 1299  
Db 1144 LSANSEPASDDPASSEMSPSTQLLFYETSASFSTEVLLQPSFQASDVDTLLKTVLPVP 1203  
Qy 1300 RDALTVPNP-----YNNRKSEGRIVYGLRPGRSYQFNVKTVSGSWKTYSK 1345  
Db 1204 SDPILVETPKVDKISITMLHLIVSNSASSENML-----HSTSVPVFVDSPTS-HMHA 1255  
Qy 1346 PIFG-SVRTKPKDKIQLNLCRPNQNSTAIACSWIPPDSDPDFGYSIECKMDTOVE--FSRK 1402  
Db 1256 SLOGLTISYASEKEYEPVLLKSSSHQVPSLYSNDELFOFANLEINQAHPPKGRHVFATP 1315  
Qy 1403 LEKEKSLNMM--LVPHKRYLVSIVKQSAQMT---SEVEDSTITWIDRPPPPPHIR 1456  
Db 1316 VLISIDPLNTLKNLHSDIELTSTKSSVTGKVFAGIPVASTFTFVSDHVSFIGNHVA 1375  
Qy 1457 VN-----EKDVLISKSNFNTVNCWFSFDTNGAVKYFTVVW-----READGDELKPEQ 1505  
Db 1376 ITAVSPHRDGSVTSTKLLFPESKAT--SELSHSAKSDAGLVGGGEDGDTDDGDDDDDRDS 1433  
Qy 1506 Q-----HPLPSYLEYR-----HNASTRYQNTYFASKCAENPNSNSKFNKLG 1549  
Db 1434 DGLSIHKMCSCSYRESQKVMNDSPTHENSL-MQNNPISYSLSENSEBENRVTSSSD 1492



QY 1550 AE--MESLGGK--RDPTQOKFCDG-----PLKPHYAYRISIRAFQOLFDE 1591  
Db 1493 SQTMDRSPKSPANGSKQNDKKEENDIQTGSALLPSPES-----KANAVLTSDE 1546  
QY 1592 -----LKE---FTKPLYSDT-----PFSLP---ITTESEP 1615  
Db 1547 ESGSGQGTSDSLNENETSDPSPADTNEKADGILAAAGDSEITPGFPQSPSTSSVTSNSE 1606  
QY 1616 LF-----GAIEG-----VSAGLFLIGMLVAVVALLICKQK- 1645  
Db 1607 VFHVSEAEAGSNSSHESRIGLAEGLESEKKAVIPLIVISALTFI--CUVWLVGILIIYWRKC 1664  
QY 1646 --VSHGRERSARLSIRDRPLSVHNLGQGNKRTCPKINQOPEGHFKLQADSNYLL 1703  
Db 1665 PQTAFHLESDTSRPRVISTPTPIPPISDDVG-----AIPKHPKXHVADLRASSGF-- 1716  
QY 1704 SKEYEELKDVGRN--QSCDIAL-----LPENRGNKRNINILPYDATRVKLSNV--DDD 1752  
Db 1717 TESEFTELKEFQEVQSCVVDLGITADSSNHPDNKHKNRYINIVAYDHSRVKLAQAEKDG 1776  
QY 1753 PCSDIYNASYIPGNFRREYIVTQGPLPGTKDDPFWKQWVEONVNIWVTOCTVEKGRVKC 1812  
Db 1777 KLTDYINANYVDGYNRPKAYIAQGPLKSTAEDEFWRMIWENHVEVIMVITNLVEKGRKC 1836  
QY 1813 DHYWPADODSLYGDLLQMLSESVLPETIREP-----KICGERQLDAHLRIHFHY 1865  
Db 1837 DQWYPAD--GSEYGNFLVTKQSVQVLAITYVRNFTLNTWKIKGSKGRPSGRVVTOYHY 1895  
QY 1866 TWPDHGVPTQSLIOFVTRVDYINRSPCAGPTVVHCSAGVGRGTCTFIADRLQOLD 1925  
Db 1896 TQWPDGVPESLPLTVFKAA--YAKRH--AVGPPVVHCSAGVGRGTCTYIIVLDSMLQIQ 1953  
QY 1926 SKDSVDYIGAHDRLHRVHMVQTECOYVYLHQCVRDVLARK 1968  
Db 1954 HEGTVNIFGFLKHRSQRNVLTQTEQYVFIHDTLVEAILSKE 1996

RESULT 14  
US-11-193-561-15  
; Sequence 15, Application US/11193561  
; Publication No. US20060024757A1  
; GENERAL INFORMATION:  
; APPLICANT: Husea, Robert  
; APPLICANT: Shorter, Simon  
; TITLE OF INVENTION: Detection of Oncofetal Fibronectin for Selection of Concepti  
; FILE REFERENCE: 17101-080001/831  
; CURRENT APPLICATION NUMBER: US/11/193,561  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: 60/592,823  
; PRIOR FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: 60/592,803  
; PRIOR FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: 60/592,825  
; PRIOR FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: 60/592,804  
; PRIOR FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: 60/592,824  
; PRIOR FILING DATE: 2004-07-30  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15  
; LENGTH: 2477  
; TYPE: PRN  
; ORGANISM: Homo Sapiens  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: GenBank NM 212482  
; DATABASE ENTRY DATE: 2005-06-10  
; US-11-193-561-15

Query Match 6.4%; Score 675.5; DB 7; Length 2477;

Best Local Similarity 20.0%; Pred. No. 1.9e-36;  
Matches 399; Conservative 321; Mismatches 769; Indels 505; Gaps 91;

QY 14 LSLQTLGLAEPERCNETLAESK-----ASSHSVSIQWRIILGSPCNF 54  
Db 682 ISIQQGHQVTRDFDTTSTSTPVTNTVGTTPSPPLVAISESVT---EITASSFVV 738  
QY 55 SLIYSDTTLGAALCPTFRID-----NTTYGCNQLQDLQAGTIY--NPKII 96  
Db 739 SWTSASDTVSG-----PRVEYELSEGBDEPYQLDLPSTATSVNI PDLPLGRKYIVNVYQI 793  
QY 97 SLDEERTVVLQTD--PLPPARFVGSKETSTGLHWVWTPSPSSKVTSEYVQLFEDNNQKI 154  
Db 794 SEDGEOSLILSTQTTAPDAPPDPTVDQDDTTSIVVRWSRPOAPITGTYRIVY----SPSV 849  
QY 155 QG--VQIQBSTSNEXETFFNLTAGSKYNIATVSGGKRSFSVY---TNGS---TVPS 204  
Db 850 EGSSTELNLPETANSVTLSDQLQGVQVNIITYAIVENQESTPVVIOQETTGTFRSDTVPS 909  
QY 205 P-----VKDIGISTKANSLISWSHSGSNBRYRLMLMDKGL--VHGQVWDKHAITSYA 256  
Db 910 PRDLQFVEVTDVKVT-----IMWTPPESAVTYRVDVIPVNLPGHGGQRLPISRNTFA 962  
QY 257 -FHGLSPGYLYNLTVMTEAAGLQNYRWKLVRTAPMEV--SNLKVNTDGLSLTKVKWQRP 314  
Db 963 EVTGLSPGVITYPKFAVSHGRESKPLTAQQTTLKLDAPTNLQFVNEDTSTVL--VRMTPPR 1021  
QY 315 GNVDSYNIT--LSHGKTIKESRVLA PWITETHEKELVPGRLYOVTVSCVSGELSAQK--- 369  
Db 1022 AQITGRVLTVGLTRGQPROQVNV--GPSVKYIPURNLOPASEYTVSLVAIKGNESPKATG 1080  
QY 370 ----MAVGTRFPDKVANLEANNRMRSLVVSWSP-----PAGMEQVYRIILFN 414  
Db 1081 VFTTLQGGSSIPP--YNTVETET---TIVITWTPAPRIGFGLGVRSQGGAPREVTS 1134  
QY 415 DSVLLNITVKEETOYVMDDTGLVPGRQYEVV--IVESNLKNKNSRCQGRTPVLA--VLQ 472  
Db 1135 SGSIWV-----SGLTPEGVYVYTIQVLRDQCDARDAPVNVKVVTPSPPTN 1179  
QY 473 LRVKHANETS--LSIMWQ---TPVAWEKYLII-----SLADRLLLIHLKSLSKDA 517  
Db 1180 LHLEANPDVTGLTVSWERSTTP--DITGYRITTTPTNGOQGNLEB-----VWH----ADQ 1229  
QY 518 KEFTFTDLVPGKYMATVTSISGLKNSSSVKGRTPV--AQVTDLHVANQMTSSSLFTNW 575  
Db 1230 SSCFTFNLSGLEYNVSVYTVKDD--KESVPISDTIPEVQPLTDLSPVDI--TDSIGLRW 1287  
QY 576 TQAGGD--VEFYQVLLI--HENVVIKNESISSETSRYSFHSKLSGSLYSV--VWTVSGGIS 631  
Db 1288 TPLNSTIIGYRITVVAAGGIPFEDFVDSVGYVTVTGLEPGIDYDISVITLINGES 1347  
QY 632 SRQVVEGRTVPSSVSGTVNNSGRNDYLSVWLVARP--DVDNYEVTLS--HDCVVQSL 688  
Db 1348 APTTLTQQTAVPPP--TDLRFTNIG--PDMRWVTWAPPSIDLTNPLVRYSPVKNEEDVAEL 1405  
QY 689 VIAKSVRECSFSSLTGRLYTVTITRSGKYENHSFS---QERTVPDKVQGVSVNSAR 744  
Db 1406 SISPSDNNAVLTNLLPGTEVVSV---SSVYEOHSTPLGRQKTGLDSTGIDFSITA 1462  
QY 745 SDYLRVSWVHATGDFDHYEVTIKKNKFI---OTKSIPKSENECVFVQLVPGRLYSVTVT 801  
Db 1463 NSF--TVHWIAPRATITGYRIR--HPEHFSGRPREDRVPHSRNSITLTNLTGTEYVVSIV 1520  
QY 802 TKSGQYEAENEQ--GNRTIPEPVKDLILNRSTEDLHVW-----SCANG 844  
Db 1521 ALNGREESPLLICQOSTVSDVPRDLEVAATPTSLISWDAPAVTVRYRITYTGETCGNS 1580  
QY 845 DVQOYEIQLLFNDMKVFPFPHLVNTATEYRFTSLTPGRQYKILVLTISGVDVQOQSAFTEG 904  
Db 1581 PQOEFVTP-----GSKSTAT---ISGLKFGVDYTIIVYATVGTGDSPPASSKPI 1625  
QY 905 TV-----PSAVKNIHISPGNATDSLTWNVTPGGGDVDSYTVSAFRHSQKVDSTQIPK 956  
Db 1626 SINRYEIDKPSQMQVTDVQDN---SISVKMLPSSSPVTGVRVT-----TTPK 1670  
QY 957 HVPEHTHRLAEAGEQVQIMIASVSGSLKNQINVVGRTPVSVQGVQVADNAYSSY----- 1010

Db 1671 NGPEPTKTAGDQTEMTIEGLQPTVEYVSVYAQN-PSGESQPLVQTAVTINIDRPKGL 1729  
QY 1011 -----SLIVSOKAAGVAERYDILLT-ENGILLRNTSEPAATTKQHKFEDLTGPKKY 1061  
Db 1730 APTDVEDSIAWESQGVSRVRYTSSPEDIHFLPAPDGEEDTASLQGLRPOSEY 1789  
QY 1062 KIQILTSSGGLFSKE-AQTEGRTPVAAVTDLRITENSTRHLSFRWTASEGELSINIFLY 1120  
Db 1790 TVSVVALHDDMESQPLIGTQSTAPAP-TDLKFTQVTPPTSLSAQWTPPNVQLTGIRVRVT 1848  
QY 1121 NPD--GNLQRAQVDPVLVQSFQNLQGRMYKMWIVTHSGELSNESFIFGRTPVASVSH 1178  
Db 1849 PKERTGPMKE-INLAPDSSSVVWVGLMVKYEVSVYALKDILTLSRPAQGVWTTLENVSP 1907  
QY 1179 LRGSNRN--TTDSLWNFWSPASGDFDYELLINPNG-TKENWKKOLTEWFOGLVPG 1235  
Db 1908 PRRARVTDATETITISWRKTETITGQVDVAVPANGQTPIORTKPDVRSYITITGLQPG 1967  
QY 1236 RKTVLWVVT-HSGDLSKNVTAERTAPSPPLMSFADIANSTLAIWK----- 1282  
Db 1968 TDYKIVYLTNDNARSPPVIDASTAIDAPSNLFLATTNSLLVSNQPPRARITGVIIK 2027  
QY 1283 -----GPPDWTVDNDFELQMLPRDALTVFNPYNNRSEGRIVYGLRPRSQQNVKTVSG 1337  
Db 2028 YEKPGSPPR-----EVPFRPRPGVT-----EATIT-GLPEPGTEYTIYVIALKN 2069  
QY 1338 DSMKTSKPIFGSVRTPKDKIQNLHCRPONSTALACSWIPDPDSDFOGYSIECRKMDTOEV 1397  
Db 2070 NQ---KSEPLIG--RKTDLSL-----PQVLTL-----PHNLHGPEILDVPSTVQKT 2111  
QY 1398 EFSRLEKEKESLLNIMLPHKRYLSIKVQSAGMTSE-----VVED-----S 1440  
Db 2112 PF-----VTHPGYDTGNGIQLPGTSQQPSVGOQMLFEHGFRRITPPT 2155  
QY 1441 TITMI-DRPPPPPHI-----RVNEKOVLSKSSINFVNCWFSFSDYNGAVKYFTVVV 1492  
Db 2156 TATPIRHRPRPYPPNVGEEIQIGHIPREDV----- 2185  
QY 1493 READGDELKPEQOHPPLSVLEYVRHNASIRVYQNTYFASKCAENPNNSNKSFNILGAEM 1552  
Db 2186 -----DYHLPHPGPL-----NPNAST----- 2202  
QY 1553 ESLGGRDPTQKQCDGPKLPHYATYRISRAFTQLFDEDLKEFTKPLYSDFPSLPITTE 1612  
Db 2203 -----GOEALSQTTISWAPQDTSEYIISCH-----PVGID 2233  
QY 1613 SEPLFGAIEGVSAGLFLIGML-----VAVVAL-----LICROKVSHGRERPS 1654  
Db 2234 EEPLQFRVPGTSTSATLTGLTRGATYNIIVEALKDQQRHKVRBEVTVGNSVNEGLNQPT 2293  
QY 1655 ARLSIRDRPLSV-HLNLGQGNKRTSCPIK-----INQEGHFMKLQADS----- 1699  
Db 2294 DSCF---DPYTVSHVAVGEWERMSESGFKLLCQCLGFGSGHF-----RCDSRWCHDNGV 2347  
QY 1700 NYLLSKEYBELKDVGRNQSCDIALPENRGKRNINILPYDATRVKLSNVDDPCSDYIN 1759  
Db 2348 NYKIGEKWDRQENQWMSG--TCLNGKGEFKCD---PHEAT-----CYD-DG 2390  
QY 1760 ASYIPGNFRREYI 1773  
Db 2391 KTYHVGEQWQKEYL 2404

RESULT 15

US-11-193-771-15  
; Sequence 15, Application US/11193771  
; Publication No. US20060024722A1  
; GENERAL INFORMATION:  
; APPLICANT: Fisher-Colbrrie, Mark  
; APPLICANT: Hickok, Durlin  
; APPLICANT: Lapointe, Jerome P.  
; TITLE OF INVENTION: Samples for Detection of Oncofetal Fibronectin and uses thereof

; FILE REFERENCE: 17101-027001/828  
; CURRENT APPLICATION NUMBER: US/11/193,771  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: 60/592,823  
; PRIOR FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: 60/592,803  
; PRIOR FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: 60/592,825  
; PRIOR FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: 60/592,804  
; PRIOR FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: 60/592,824  
; PRIOR FILING DATE: 2004-07-30  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15  
; LENGTH: 2477  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: GenBank NM 212482  
; DATABASE ENTRY DATE: 2005-06-10  
US-11-193-771-15

Query Match 6.4%; Score 675.5; DB 7; Length 2477;  
Best Local Similarity 20.0%; Pred. No. 1.9e-36;  
Matches 399; Conservative 321; Mismatches 769; Indels 505; Gaps 91;  
QY 14 LSLQLGLABPERCNTTLAESK-----ASSHSVSTQWRILGSPCNF 54  
Db 682 ISIQOYGHQEVTRFDPTTSTSTPTVTNTGTTPSPLVATSESVT---EITASSFVV 738  
QY 55 SLIYSSDTLGAALCPTRFD-----NTTYGCLQDLQAGTIV--NFKII 96  
Db 739 SWSASDTVSG-----FRVEYELSEGEDPEQYLDLSTATSVNIPDLPLGRKIVVYQI 793  
QY 97 SLDEERTVVLQTD--PLPPAREFVSKKETTGLHVMWTFSSGKVTSYEVQLFDENNQKI 154  
Db 794 SEDGEQSLILSTQTTAPDAPPDPTVDQVDDTSIVVWMSRPPQAPITGRIYV---SPSV 849  
QY 155 QG--VOIQBSTSWNETFFNLTAGSKYNIATAVSGGKESFSVY-----TNGS----TVPS 204  
Db 850 EGSSTELNLPETANSVTLSDLPQGVQYNTIYVAEENQESTPVVIOQETGTGTPRSDTVPS 909  
QY 205 P-----VKDIGISTKANSLLISWHSNGVRYRLMLMDKGL--VHGVVDKHAATSYA 256  
Db 910 PRDLQFVEVTDVKVI-----IMTTPPESAVTGHYVDVIVPNLPGHGORLPISNTFA 962  
QY 257 -FHGLSPGYLYNLTVNTEAAGLQNYRWKLVRTAPMEV--SNLKVTNDGSLTSLKVKWQRP 314  
Db 963 EVTGLSPGYTYFYKFAVSHGRESKPLTAQQTTKLDAPTNLQFVNETDSTVL--VRWTPPR 1021  
QY 315 GNWDSYNIIT--LSHGKTIKESVILAPWITETHFKELVORLYQVTVSCVSGELSAOK--- 369  
Db 1022 AQITGYRLTVGLTRGQPRQYV--GPSVSKYPLRLNLPQASEYTVSLVAKNQESPKATG 1080  
QY 370 ----NAVGTFFDKVANLEANNRMRSLVWSWP-----PAGWQEYRILLEN 414  
Db 1081 VFTTLPQSGSIPP--YNTVETET---TIVITWTAPRIFGLGVRPSSQGGEPREVTSD 1134  
QY 415 DSVVLNITVGBEETQYVMDDTGLVPGROYEVEV--IVESGNLKNSERCQRTVPLA-VLQ 472  
Db 1135 SGSIVV-----SGLTPGVEYVYTIQVLRDQGERDAPVKNKVVTFPSPTN 1179  
QY 473 LRVKGANETS-LSIMWQ-----TPVAWEKYLII-----SLADRLILLIHKLSKDA 517  
Db 1180 LHLEANPDGTGLTVSWEKSTTP--DITGYRITTTTPTNGQQGNSLEB-----VH- 1229  
QY 518 KEFTFTDLVPRGKYMATVTSISGDLKNSSSVKGRTPV--AQVTDLHVANQGMSTSLFTNW 575  
Db 1230 SSCFTDNLSFGLYNNVSVVTKDD--KESVPISDITIPEVQTLTDLISFVDI--TDSGIGLRW 1287  
QY 576 TQAQGD-VFVYQVLLI--HENYVIKNESISSETSRYSFHSLSKSGSLYSV-VVTVSGGIS 631

Db 1288 TPLNSSIIGRIITVAAAGGIPFEDFVSSVCYVTVGLPGCIDYDVSITLINGES 1347  
Qy 632 SRQVVEGRTPVSSVGVTVNNNSGRNDYLSVSMVLVAPG-DVDNVEVTVLS--HDKVQVQSL 688  
Db 1348 APTTLTQQTAVPPD-TDLRFTNIG-POTMRVTWAPPDSIDLTFELVRSVPKNEEDVAEL 1405  
Qy 689 VIAKSVRECSFSLTPCGLTYVTITTSKGYENHSFS---QERTVPDKVQGVSVNSAR 744  
Db 1406 SISPSDNVAVLNLPLPTEYVSV---SSVYQESFPLRGROKTGLDSTGIDFSDITA 1462  
Qy 745 SDYLRVSWHATGDFHYEVTIKNNFI---QTSIPKSENECVFVQLVPGRLYSVTVT 801  
Db 1463 NSF-TVMIAPRATITGYRIR-HPEHFSGRPRDRVPHSRNSITLNLTPGTEYVVSIV 1520  
Qy 802 TKSQYEAHQ-NGRTPIPBVKDLTLNRSTEDLHVTW-----SGANG 844  
Db 1521 ALNGREESPLLIGQOSTVSDVPRDLVVAAPTSLISWDAPAVTVRYRYITYGETGNS 1580  
Qy 845 DVQOYEIQLFNDMKVPPPHLVNTATEYRFTSLTPGQYKILVLTISGDVQOSAFIEGF 904  
Db 1581 PVQEFVTP-----GSKSTAT---ISGLKPGVDYITTVYAVTGRGDSPASSKPI 1625  
Qy 905 TV-----PSAVKNIHSPNGATDSLTVNWTGPGGDVDSYTVSAFRHSQKVDSDOTIPK 956  
Db 1626 SINVRTIDKPSQOVTDVQDN-----SISVKWLPSSSPVTGYRV-----TTPK 1670  
Qy 957 HVFEHTFHRLEAGEQYQIMIASVSGSLKNOINVVGRTPVPSVQGVIAADNAYSSY----- 1010  
Db 1671 NGPGTKTKTAGPDQTEMTIEGLQPTVEYVSVYAQN-PSGESQPLVQTAVTNIDRPKGL 1729  
Qy 1011 -----SLIVSWKKAAGVAERYDILLT-ENGILLRNTSEBAPTQKHFPEDLTPGKXY 1061  
Db 1730 APTDVEDDSIKIAWESQGVSRVTVYSSPEDGIELFPAPDGEEDTAELOGLRPGSEY 1789  
Qy 1062 KIOILTVSGGLFSKE-AQTEGRTVPAAVTDLRIENSTRHLSFRWATASEGELSWYNIFYL 1120  
Db 1790 TVSVVALHDMESQPLIGTOSTAIPAP-TDLKFTQVTPTSLSAGTAPPNVQLTGYRVVT 1848  
Qy 1121 NPD--GNLORAAQVDPLVQFSFQNLQGRMYKVIIVTHSGELSNESFIFGRTPVASVSH 1178  
Db 1849 PKEKTGPMKE-INLAPDSSVWSGLMVATKYEVSVALKDLTSLRPAQGVVITLNVSP 1907  
Qy 1179 LRGSNRN--TTDSLWFNWSASGDFDYELLNPNNG-TKENWKOXDLTEWRPQGLVPG 1235  
Db 1908 PRARVTDATETITISWRKTETITIGFQVDAVPANGQTPQRTIKPDVRSYITITGLQPG 1967  
Qy 1236 RYVLMWVT-HSGDLSNKTAESESTAPSPPSLMSFADIANSTSLAITWK----- 1282  
Db 1968 TDYKIYDITLNDNARSFVIDASTAIDAPSNRFLATTNPSNLVSWQPPRARTGYIHK 2027  
Qy 1283 -----GPPDWTVDNDFELOWLPRDALTVFNPNYNNKSEGRIVYGLRPGSYQFNVKTVSG 1337  
Db 2028 YEKPGSPPR-----EVVPRPRGVT-----EATIT-GLEPGTEYTIIVIALKN 2069  
Qy 1338 DSWKTYSKPIFGSVRTKPKIQNLHCRPQNSTAIACSWIPDSDPDGYSIECRMDTQEV 1397  
Db 2070 NQ---KSEPLIG-RKKTDEL-----POLVTL-----PHNHLHGPEILDVPSTVQKT 2111  
Qy 1398 EFSRLEKEKESLLNIMLVPHKRYLVSVIKVQSAGMTSE-----VVED-----S 1440  
Db 2112 PF-----VTHPGYDTGNGIQLPGTSGQPSVQGOQMIPEHGFRRTPPT 2155  
Qy 1441 TITMI-DRPPPPPPPHI-----RVNEKDVLSKSSINFVNCWSWFSFDTNGAVKYFTVVV 1492  
Db 2156 TATFIRHRPRPPYPPNVGEEIQIGHIPREDV----- 2185  
Qy 1493 READGSDLPKQOHPPLPSYLEYRHNASIRVYQTNYPASKAENPNNSKSNFIKLGAE 1552  
Db 2186 -----DYHLPHGPGGL-----NPNAST----- 2202  
Qy 1553 ESIGKEDPTQOKFCDCGFLPKHTAYRISIAFTQLFDEDLKEFTKPLYSDFPSLPITTE 1612

Db 2203 -----GOEALSQTTISWAPPQDTSYIISCH-----PVGTD 2233  
Qy 1613 SEPLFGAIEGVSAGLFLIGML-----VAVVAL-----LICROKVSUHGRRPS 1654  
Db 2234 EEPLOFRVPGTSTSATLTGLTRGATYNIIVEALKDQQRHKVRBEVTVVGNVNEGLNOPT 2293  
Qy 1655 ARLSIRDRPLSV-HLNLGOKGNKRTSCPIK-----INQFEGHFMKLOADS----- 1699  
Db 2294 DDSCF---DPYTVSHVAVGDEWERMESGFKLLCQCLGFGSGHF---RCDSSRWCHONGV 2347  
Qy 1700 NYLLSKEYEELKDVGRNQSCDIALLPENRGNKRYNNILPYDATRVKLSNVDDDDPCSDYIN 1759  
Db 2348 NYKIGEKWDRQGENGQWMSK--TCLNGKGEFKCD---PHEAT-----CYD-DG 2390  
Qy 1760 ASYIPGNFRREYI 1773  
Db 2391 KTYHVGEOQWOKBYL 2404

Search completed: March 10, 2006, 19:07:52  
Job time : 50.3095 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2006, 18:27:04 ; Search time 48.7073 Seconds  
(without alignments)  
3944.887 Million cell updates/sec

Title: US-10-633-742-2  
Perfect score: 10483  
Sequence: 1 MLSHGAGLAWITLSLLQTG.....PPIYENVPEYHRDPVYSRH 1997

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR.80:\*  
1: Pirl.\*  
2: Pirl2.\*  
3: Pirl3.\*  
4: Pirl4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10483	100.0	1997	1 S12050	protein-tyrosine-p
2	2807	26.8	583	2 S17671	protein-tyrosine-p
3	1530	14.6	1557	2 D41214	protein-tyrosine-p
4	1526	14.6	1630	2 C41214	protein-tyrosine-p
5	1475.5	14.1	1615	2 B49502	protein-tyrosine-p
6	1475.5	14.1	1767	2 A49502	protein-tyrosine-p
7	1249	11.9	1337	1 I38670	protein-tyrosine-p
8	1243.5	11.9	1238	2 S68700	HPTP beta-like tyr
9	1198	11.4	1711	1 A55148	protein-tyrosine-p
10	1097	10.5	2302	2 T14328	protein-tyrosine-p
11	939	9.0	1367	2 T21913	hypothetical prote
12	913.5	8.7	1188	1 A57084	protein-tyrosine-p
13	908.5	8.7	1187	1 A53661	protein-tyrosine-p
14	906.5	8.6	1216	2 S60613	protein-tyrosine-p
15	886	8.5	1118	1 A49724	protein-tyrosine-p
16	877	8.4	1226	2 J07503	protein-tyrosine-p
17	836	8.0	1898	2 S46216	leukocyte antigen-
18	832.5	7.9	1912	2 A56178	protein-tyrosine-p
19	832	7.9	1894	2 A54689	protein-tyrosine-p
20	829	7.9	1691	1 D54689	protein-tyrosine-p
21	826	7.9	1897	1 TDHULK	leukocyte antigen-
22	799.5	7.6	2029	1 TDFFLK	protein-tyrosine-p
23	780.5	7.4	405	2 I49372	protein-tyrosine-p
24	778.5	7.4	1907	2 S50893	protein-tyrosine-p
25	777	7.4	405	2 S68250	protein-tyrosine-p
26	776	7.4	2051	2 T30938	receptor tyrosine
27	768	7.3	1290	2 A56493	leucocyte common a
28	763.5	7.3	1863	2 S46217	protein-tyrosine-p
29	748.5	7.1	1501	2 I58148	protein-tyrosine-p

30	747.5	7.1	1499	2 I50212	protein-tyrosine-p
31	744.5	7.1	1262	1 B48758	protein-tyrosine-p
32	744.5	7.1	1496	1 A48758	protein-tyrosine-p
33	724	6.9	1437	2 T31093	probable protein-t
34	719.5	6.9	1452	1 S17670	protein-tyrosine-p
35	718.5	6.9	1452	1 S17669	protein-tyrosine-p
36	713.5	6.8	1585	2 T19121	probable protein-t
37	711	6.8	1440	2 JC6312	protein-tyrosine-p
38	696	6.6	1457	1 A48066	protein-tyrosine-p
39	687.5	6.6	2314	1 A46151	protein-tyrosine-p
40	674	6.4	4006	2 T09070	probable tenascin
41	672	6.4	1445	1 A48148	protein-tyrosine-p
42	668	6.4	2477	2 S14428	fibronectin precur
43	661.5	6.3	1301	1 A41622	protein-tyrosine-p
44	657	6.3	1273	1 TDRILT	leukocyte common a
45	655	6.2	1422	2 T42636	protein-tyrosine-p

## ALIGNMENTS

## RESULT 1

S12050  
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type beta precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: S12050; S15818; S15819  
R:Krueger, N.X.; Streuli, M.; Saito, H.  
EMBO J. 9, 3241-3252, 1990  
A:Title: Structural diversity and evolution of human receptor-like protein tyrosine phos  
A:Reference number: S12049; MUID:91006018; PMID:2170109  
A:Accession: S12050  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1997 <XRU>  
A:Cross-references: UNIPROT:P23467; UNIPARC:UPI0000034765; GB:X54131; NID:g35787; PION:C  
P;de Vries, L.; Li, R.Y.; Ragab, A.; Ragab-Thomas, J.M.F.; Chap, H.  
FEBS Lett. 282, 285-288, 1991  
A:Title: Expression of a truncated protein-tyrosine phosphatase mRNA in human lung.  
A:Reference number: S15818; MUID:91243813; PMID:1645282  
A:Accession: S15818  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1872-1911, 'VHMVLQK' <VRI>  
A:Cross-references: UNIPARC:UPI0000173861  
A:Accession: S15819  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1872-1997 <VR2>  
A:Cross-references: UNIPARC:UPI0000173861  
C:Genetics:  
A:Gene: GDB:PTPRB; PTPB  
A:Cross-references: GDB:127352; OMIM:176882  
A:Map position: 12q15-12q21  
C:Superfamily: protein-tyrosine-phosphatase, receptor type beta; fibronectin type III re  
C:Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; recept  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-1997/Product: protein-tyrosine-phosphatase, receptor type beta #status predicted <M  
F:23-1625/Domain: extracellular #status predicted <EXT>  
F:1626-1642/Domain: transmembrane #status predicted <TMN>  
F:1643-1997/Domain: intracellular #status predicted <INT>  
F:1727-1952/Domain: protein-tyrosine-phosphatase homology <PTP>  
F:1904/Active site: Cys (phosphocysteine intermediate) #status predicted  
F:1910/Binding site: substrate phosphate (Arg) #status predicted

Query Match 100.0%; Score 10483; DB 1; Length 1997;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1997; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLSHGAGLAWITLSLLQTLAEPNCNFTLAESKASHSVSIQWRILGSPCNFSLIYSS 60

Db 1 MLSHGAGLAWITLSLLQTLAEPNCNFTLAESKASHSVSIQWRILGSPCNFSLIYSS 60

Qy 61 DTLGAALCPTPRIINTTYGCMQLDQAGTIYNFKIISLDEERTVVLQTDPLPPARFGVSK 120  
Db 61 DTLGAALCPTPRIINTTYGCMQLDQAGTIYNFKIISLDEERTVVLQTDPLPPARFGVSK 120  
Qy 121 EKTTSGLHWMTPTSSGKVTSEYVOLFDENNOKIQGVQIQBSTSWNEYTFNLTAGSKYN 180  
Db 121 EKTTSGLHWMTPTSSGKVTSEYVOLFDENNOKIQGVQIQBSTSWNEYTFNLTAGSKYN 180  
Qy 181 IAITAVSGGKRSFVSFYNGSTVPSPVKDIGISTKANSLIISWSHSGSNVERYRLMLMDKG 240  
Db 181 IAITAVSGGKRSFVSFYNGSTVPSPVKDIGISTKANSLIISWSHSGSNVERYRLMLMDKG 240  
Qy 241 ILVHGGVVDKATSYAFHGLSPGYLYNLTVMTAEAGLQNYRWKLVRTAPMEVSNLKVTDND 300  
Db 241 ILVHGGVVDKATSYAFHGLSPGYLYNLTVMTAEAGLQNYRWKLVRTAPMEVSNLKVTDND 300  
Qy 301 GSLTSLVKWQRPNGNDVSNITLISHKGTIKESRYLAPWITETHFKELVPGRLYQVTVSC 360  
Db 301 GSLTSLVKWQRPNGNDVSNITLISHKGTIKESRYLAPWITETHFKELVPGRLYQVTVSC 360  
Qy 361 VSGELSAQKMAVGRTFPDKVANLEANNRMRSLLVSVSPAGDWEQYRILLFNDSVULL 420  
Db 361 VSGELSAQKMAVGRTFPDKVANLEANNRMRSLLVSVSPAGDWEQYRILLFNDSVULL 420  
Qy 421 NITVGEETOYVMDDTGLVPGRQYEVEVIVESGNLKNSERCQGRTPVLAVLQLRVKHANE 480  
Db 421 NITVGEETOYVMDDTGLVPGRQYEVEVIVESGNLKNSERCQGRTPVLAVLQLRVKHANE 480  
Qy 481 TSLSIMQWTPVAEWEKIIISLADRDLLIHKLSLSDAKETFTDLVPGRYMATVTSISG 540  
Db 481 TSLSIMQWTPVAEWEKIIISLADRDLLIHKLSLSDAKETFTDLVPGRYMATVTSISG 540  
Qy 541 DLKNSSVKGRTPVAQVTDLHVANQGMTSSLFTNWTQAGDVEFYQVLLIHENVVIKNES 600  
Db 541 DLKNSSVKGRTPVAQVTDLHVANQGMTSSLFTNWTQAGDVEFYQVLLIHENVVIKNES 600  
Qy 601 ISSETSRYSFHLKSGSLYSVVVTVSGGSSRQVVVEGRTVPSSVSGVTVNNSGRNDYL 660  
Db 601 ISSETSRYSFHLKSGSLYSVVVTVSGGSSRQVVVEGRTVPSSVSGVTVNNSGRNDYL 660  
Qy 661 SVSHVAPGDVNDYEVTLSDHGKVVQSLVIAKSVRECSFSLTPGRLYTVTITTRSGKYE 720  
Db 661 SVSHVAPGDVNDYEVTLSDHGKVVQSLVIAKSVRECSFSLTPGRLYTVTITTRSGKYE 720  
Qy 721 NHPFSQERTVPDKQGVSVNSARS DYLRVSWHATGDFDHYEVTIKNKNFQTSKIPK 780  
Db 721 NHPFSQERTVPDKQGVSVNSARS DYLRVSWHATGDFDHYEVTIKNKNFQTSKIPK 780  
Qy 781 SENECVFQVLPGRLYSVTVTTKSGQYEAENGNGRTIPEPVKDLTLNRSTEDLHVTVS 840  
Db 781 SENECVFQVLPGRLYSVTVTTKSGQYEAENGNGRTIPEPVKDLTLNRSTEDLHVTVS 840  
Qy 841 GANGDVQDYEIQLLFNDMKVPPPHLUNTAETEFTSLTPGRQYKILVLITISGDVQOSAF 900  
Db 841 GANGDVQDYEIQLLFNDMKVPPPHLUNTAETEFTSLTPGRQYKILVLITISGDVQOSAF 900  
Qy 901 IEGFTVPSAVKNLHI SPNGATDSLTVNWTGGGDVDSYTVSAPRHSOKVDQSOTIPKHVE 960  
Db 901 IEGFTVPSAVKNLHI SPNGATDSLTVNWTGGGDVDSYTVSAPRHSOKVDQSOTIPKHVE 960  
Qy 961 HTFHLRLEAGQYQIMIASVSGSLKNQINNVGRTVPASVQGVVIADNAYSSSLIVSMQKAA 1020  
Db 961 HTFHLRLEAGQYQIMIASVSGSLKNQINNVGRTVPASVQGVVIADNAYSSSLIVSMQKAA 1020  
Qy 1021 GVAERYDILLTENGILLRNTSEBATTQKHFDLTPGKKYKIQILTVSGGLSKERQTE 1080  
Db 1021 GVAERYDILLTENGILLRNTSEBATTQKHFDLTPGKKYKIQILTVSGGLSKERQTE 1080  
Qy 1081 GRTVPAAVTDLRITENSTRHLSFRWTASEGELSWYNIFLYNPDGNLQERQVDPVQOSFS 1140  
Db 1081 GRTVPAAVTDLRITENSTRHLSFRWTASEGELSWYNIFLYNPDGNLQERQVDPVQOSFS 1140  
Qy 1141 FONLLQGRMYKMWIVTHSGELSNBSFIFGRTVPASVSHLRGSNRNTTDSLWFNWPASGD 1200

Db 1141 FONLLQGRMYKMWIVTHSGELSNBSFIFGRTVPASVSHLRGSNRNTTDSLWFNWPASGD 1200  
Qy 1201 FDFVELILYNPNGTKKENWKDKDLTEWRFOGLVPGRYKVLVWVTHSGDLNKNVTAERTTA 1260  
Db 1201 FDFVELILYNPNGTKKENWKDKDLTEWRFOGLVPGRYKVLVWVTHSGDLNKNVTAERTTA 1260  
Qy 1261 PSPSLMSFADIANSTSLAITWKGPPDWTNDYDFELQWLPDALTVPFNPNRKRSEGRIVY 1320  
Db 1261 PSPSLMSFADIANSTSLAITWKGPPDWTNDYDFELQWLPDALTVPFNPNRKRSEGRIVY 1320  
Qy 1321 GLRPGRSYQNVKTVSGDSWKTSKPIFGSVRTKPKIQNLHCRPQNSTAIACSWIPPPDS 1380  
Db 1321 GLRPGRSYQNVKTVSGDSWKTSKPIFGSVRTKPKIQNLHCRPQNSTAIACSWIPPPDS 1380  
Qy 1381 DFDGYSIECKMDTOEVEFSRKLKESKLLINIMWLPHKRYLVSIKVQSGAGMTSEVVEDS 1440  
Db 1381 DFDGYSIECKMDTOEVEFSRKLKESKLLINIMWLPHKRYLVSIKVQSGAGMTSEVVEDS 1440  
Qy 1441 TITMIDRPPPPPHIRVNEKDVILISKSSINFTVNCWFSDTNGAVKYFTVVVREADGSDE 1500  
Db 1441 TITMIDRPPPPPHIRVNEKDVILISKSSINFTVNCWFSDTNGAVKYFTVVVREADGSDE 1500  
Qy 1501 LKPEQOHPSPSYLEYRHNASIRVYQNTYFASKCAENPNNSKSNFKIKLGAEMESLGGKRD 1560  
Db 1501 LKPEQOHPSPSYLEYRHNASIRVYQNTYFASKCAENPNNSKSNFKIKLGAEMESLGGKRD 1560  
Qy 1561 PTQCKECGDLKPHYATYRISIRAFQOLFDDELKEFTKPLYSDFPSLPITTESEPLFGAI 1620  
Db 1561 PTQCKECGDLKPHYATYRISIRAFQOLFDDELKEFTKPLYSDFPSLPITTESEPLFGAI 1620  
Qy 1621 EGVSAGLFLIGMLVAVVALLICRQKVSIGRERSARLSIRDRPLSVHNLGOKGNRKT 1680  
Db 1621 EGVSAGLFLIGMLVAVVALLICRQKVSIGRERSARLSIRDRPLSVHNLGOKGNRKT 1680  
Qy 1681 CPIKINQFEGHFMKLQADSNLYLSKEYEELKDVGRNOSCDIALLENRGNKRYNNILPYD 1740  
Db 1681 CPIKINQFEGHFMKLQADSNLYLSKEYEELKDVGRNOSCDIALLENRGNKRYNNILPYD 1740  
Qy 1741 ATRVKNLSDVDDPCSDYINASYIPGNFRREYIVTQGLPGTKDDFKMWVQONVNIIVM 1800  
Db 1741 ATRVKNLSDVDDPCSDYINASYIPGNFRREYIVTQGLPGTKDDFKMWVQONVNIIVM 1800  
Qy 1801 VTQCEKGRVKCDHYVPADQDSLYYGDLLIQLMSESVLPETWITREFKICGEEQLDAHRLI 1860  
Db 1801 VTQCEKGRVKCDHYVPADQDSLYYGDLLIQLMSESVLPETWITREFKICGEEQLDAHRLI 1860  
Qy 1861 RHPHYTVWPDHGVPTTQSLIQFVTRVRYINRSPGAGPTVHCSAGVGRGTGTFIALDRI 1920  
Db 1861 RHPHYTVWPDHGVPTTQSLIQFVTRVRYINRSPGAGPTVHCSAGVGRGTGTFIALDRI 1920  
Qy 1921 LQOLDKSDSDYIYGAVHDLRLHVRHMVOTECQVYVYLHQCVRDYLARKLRSEQENPLFPI 1980  
Db 1921 LQOLDKSDSDYIYGAVHDLRLHVRHMVOTECQVYVYLHQCVRDYLARKLRSEQENPLFPI 1980  
Qy 1981 YENVNPEYHRDPVYSRH 1997  
Db 1981 YENVNPEYHRDPVYSRH 1997

## RESULT 2

S17671

protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type beta - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 22-Nov-1993 #sequence\_revision 15-Mar-1996 #text\_change 31-Dec-2004

R;Accession: S17671; S40287

R;Gabbink, M.F.B.G.; van Eeten, I.; Hateboer, G.; Suijkerbuijk, R.; Beijersbergen, R.L.;

FEBS Lett. 290, 123-130, 1991

A;Title: Cloning, expression and chromosomal localization of a new putative receptor-like

A;Reference number: S17669, MUID:92008644; PMID:1655529

A;Accession: S17671

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: mRNA

A;Residues: 1-583 <GB>  
A;Cross-references: UNIPROT:Q64497; UNIPARC:UPI000017705E; EMBL:X58289  
R;Hendriks, W.; Brugman, C.; Zeeuwen, P.; Schepens, J.; Wieringa, B.  
submitted to the EMBL Data Library, June 1993  
A;Description: Assessment of the expression levels of murine protein-tyrosine phosphatase  
A;Reference number: 540280  
A;Accession: 540287  
A;Molecule type: mRNA  
A;Residues: 377-483, 'T', 485-486 <HEN>  
A;Cross-references: UNIPARC:UPI0000021846; EMBL:Z23056; NID:9438149; PIDN:CAA80591.1; PI  
C;Superfamily: fibronectin type III repeat homology; protein-tyrosine-phosphatase homolo  
C;Keywords: phosphoprotein; phosphoric monoester hydrolase; receptor; transmembrane prot  
F;311-536/Domain: protein-tyrosine-phosphatase homology <PTP>  
F;488/Active site: Cys (phosphocysteine intermediate) #status predicted  
F;494/Binding site: substrate phosphate (Arg) #status predicted

Query Match 26.8%; Score 2807; DB 2; Length 583;  
Best Local Similarity 92.8%; Pred. No. 1e-148;  
Matches 526; Conservative 14; Mismatches 27; Indels 0; Gaps 0;

Qy 1419 KRYLVSIKVSAGMTSEVEDSTTMDRPPPPPHIRVNEKOVLSKSSINFVNCSEWF 1478  
Db 3 KRYLVSIKVSAGMTSEVEDSTTMDRPPPPPHIRVNEKOVLSKSSINFVNCSEWF 62

Qy 1479 SDTNGAVKYFTVVREADGSDPEQOHPPLPSYLEVRHNASIRVOTNYFASKCAENPN 1538  
Db 63 SDTNGAVGYFAVVREADSDMDPEQOHPPLPSYLEVRHNASIRVOTNYFASKCAESP 122

Qy 1539 SNSKSFNIKGAEMESLGGKRDPTQOKFCDGPLKPTAYRISIRAFQFDEDLKFTKP 1598  
Db 123 SSSKSFNIKGAEMDSLGGKCDPQOKFCDGPLKPTAYRISIRAFQFDEDLKFTKP 182

Qy 1599 LYSTDFSLPTTSEPLFGAIEGVSAFLIGMLVAVALLICROKSHGRERPSARLS 1658  
Db 183 LYSTDFSMPTTSEPLFGVIEGVSAFLIGMLVALVAFICROKASHRERPSARLS 242

Qy 1659 IRRDRPLSVHLNCGKGRKTCPIKINOPGHEFMKQADSNYLLSKEYELKDVGKNS 1718  
Db 243 IRRDRPLSVHLNCGKGRKTCPIKINOPGHEFMKQADSNYLLSKEYELKDVGKRSQS 302

Qy 1719 CDIALLENKRNKRNINILPYDARVKLSNVDDPCSDYINASYPGNFRREYIVTQGP 1778  
Db 303 CDIALLENKRNKRNINILPYDASRVKLSNVDDPCSDYINASYPGNFRREYIVTQGP 362

Qy 1779 LPGTKDDFWKMWQONVHNIWMVTCVEKGRVKCDHYWPDADQDSLYYGDILQMLSSVL 1838  
Db 363 LPGTKDDFWKMWQONVHNIWMVTCVEKGRVKCDHYWPDADQDPLYYGDLILQWSSVL 422

Qy 1839 PEWTIREFKICGEQOLDARLIRHFHYTWPDHGVPTTQSLIQFVRTVDYINRSPGAG 1898  
Db 423 PEWTIREFKICGEQOLDARLIRHFHYTWPDHGVPTTQSLIQFVRTVDYINRSPGAG 482

Qy 1899 PTVVHCAGVGRCTFTALDRILQOLDSKSDVIYGAVHDLRLHRVMVOTECQVYVYHQ 1958  
Db 483 PSVVHCAGVGRCTFTVALDRILQOLDSKSDVIYGAVHDLRLHRVMVOTECQVYVYHQ 542

Qy 1959 CVRDVLARKLRSEQENPLPPIYENVN 1985  
Db 543 CVRDVSEQRNCGNEQKGGVSIYENVN 569

RESULT 3  
D41214  
protein-tyrosine-phosphatase (EC 3.1.1.3.48), receptor type 10D, short splice form precure  
C;Species: Drosophila melanogaster  
C;Date: 28-May-1992 #sequence\_revision 28-May-1992 #text\_change 09-Jul-2004  
C;Accession: D41214; A41215  
R;Yang, X.; Seow, K.T.; Bahri, S.M.; Oon, S.H.; Chia, W.  
Cell 67, 661-673, 1991  
A;Title: Two Drosophila receptor-like tyrosine phosphatase genes are expressed in a subs  
A;Reference number: A41214; MUID:92034988; PMID:1657401  
A;Accession: D41214  
A;Status: nucleic acid sequence not shown; not compared with conceptual translation

A;Molecule type: mRNA  
A;Residues: 1-1557 <XAN>  
A;Cross-references: UNIPROT:P35992; UNIPARC:UPI00000177055; GB:M80465  
R;Tian, S.S.; Tsoulfas, P.; Zinn, K.  
Cell 67, 675-685, 1991  
A;Title: Three receptor-linked protein-tyrosine phosphatases are selectively expressed o  
A;Reference number: A41215; MUID:92034989; PMID:1657402  
A;Accession: A41215  
A;Molecule type: mRNA  
A;Residues: 1-904, 'L', 906-1125, 'Q', 1126-1165, 'YR', 1168-1171, 'A', 1173-1215, 'L', 1217-1456,  
A;Cross-references: UNIPARC:UPI000002B1D9; GB:M80538; NID:g158644; PIDN:AAA28952.1; PID:  
C;Genetics:  
A;Gene: FlyBase:Ptp10D  
A;Cross-references: FlyBase:FBgn0004370  
C;Superfamily: protein-tyrosine-phosphatase, receptor type 4E; fibronectin type III repe  
C;Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; recept  
F;1197-1213/Domain: transmembrane #status predicted <TMN>  
F;1214-1557/Domain: intracellular #status predicted <INT>  
F;1295-1515/Domain: protein-tyrosine-phosphatase homology <PTP1>  
F;1467/Active site: Cys (phosphocysteine intermediate) #status predicted  
F;1473/Binding site: substrate phosphate (Arg) #status predicted

Query Match 14.6%; Score 1530; DB 2; Length 1557;  
Best Local Similarity 27.6%; Pred. No. 8.2e-77;  
Matches 510; Conservative 273; Mismatches 645; Indels 418; Gaps 73;

Qy 265 LYNLTVMTEAAGLQ-----NYRW-----KLVRTAPMEVS--NUKVTNDGSLT 304  
Db 2 LYQLSKATTRILKQKAVPQHRWLSLAFLLAFTLKDVRCADLAISIPNPGLLDGA-- 59

Qy 305 SLKVKQRPNGVDSNYITLSHGKTIKESVLAIPWITETHFKELVPCRLQ----- 355  
Db 60 SYRLDYS-PPFGYPEPNTTIA-----SREIG---DEIOFSRALPGTKNFWLYNFT 108

Qy 356 -----VTVSSCVSGELSAQKMAVGRTFPDKVANLEAN--NGRMRLSLVSWSPPA--GDW 405  
Db 109 HHDWLWTVTIIT-----TAPDPSNLQVVRSG--KNAIILMSPTQGSY 151

Qy 406 EQYRILLFNDSSVLLNTVIGKEETQYVMDT-----GLVPGROYEVE-----VI 449  
Db 152 TAFKIKVLG-----LSEASSSYNRTFQVNDNTFOHSVKELTPGATYQVQATYIDGESVA 207

Qy 450 VESGNLKNSECOGRTPVLAQLRVKHANETSLSIMQOT--PVAWEKIILSI---AD 503  
Db 208 YTSRNFYTKNTPKGFI-----VMFRNETLLVWQPPYPAGIYTHYKVSIEPDPAN 259

Qy 504 RDLLIHKLSKDA--KEFTFTDLVPGKYMATVTSIGD--LKNSSSVKGRTPVPAQVTDLH 561  
Db 260 DSVLYVEKEGEPQPAQAAPKGLVPGAYNISVQTMSEDEISLPTTAQYRTVPLRLNVT 319

Qy 562 VANQGMTSSLF--TNWTOAQGDVEF--YQVLLI-----HENVVKNESISSETSRYSFSLK 614  
Db 320 FDRDFITSNFRVLWEAPKGISSEFDKYQVSVATTTRQSTVPSNEPVAFSDFR---DIAE 376

Qy 615 SGLSYVWVTVTSGGITSRRQVVEGRTPVSSVSGVTNNSGRNDYLSVSWLVAQDV--DN 673  
Db 377 PGKTFNVIVKTVSGKVTSPATGDTLRLPLFVRLNRSINDDKTNMTITWEADPASTQDE 436

Qy 674 YEVTLSHDGKVVQSLVIKSVRECSFSLTPGRLYTVTITTRSKYENHNSPQERTVPDK 733  
Db 437 YRIV-----YHELETFNGDTSITLT----- 456

Qy 734 VQGVSVNSARSYLRVSVVHATGDFDHYEVTIKKNKNFIQTSPKSENECVFVLVPG 793  
Db 457 -----DTRFTLES-----LLPG 469

Qy 794 RLYSVTVTTKSGQYEAHQ-----NGRTIPBPVKDLTLNRSTEDLHVTW--SGANDV 846  
Db 470 RNYSLSVQAVSKKESNETSIFVVTFRSPPIEDLKSIRM-----GLNISWKSVDVNSKQ 523

Qy 847 DQYEIQLLEN--DMKVFPFPHLVNTATEVRF--SLTPGRQYKILVLTISGDVQQS--A 899  
Db 524 EQYEVLSRNGTSDLR-----TQTKESRLVIKNLQPGAVELKVPANVSHDLRSPPHA 576









```
Db 690 TVSGVSPHLELVNTPQPVNVPLVDSRLTLEWPRPDGHVDFYTLKWPTDB-- 747
Qy 1127 QEAQVDPLVQ-----SFSQNLQGRMYKX-VIVTHSGELSNESFIFGRVTP--- 1173
Db 748 EDRVEFNKVTQLEDLSGPSVRIPEDLSPGQRYFEVQASNGIRSGTTHLSTRTMLPIQ 807
Qy 1174 -----ASVSHLGRNRTTDSLWNFNSPASGD---FDYELLINPNGTKEENWKDKDLT 1225
Db 808 SDVFIANAGHGOQDETIT-----LSYTPPADSTRFDIYRFSMGDPTIKOKEKLANDTER 863
Qy 1226 EWRQGLVPGKYYLVVVVTHSGDLSN-KVTAESRTAPSPGLMFGADIANSLAITWKGP 1284
Db 864 KLSFSGLTGPKLYNVVTVVSGGVASLPVQRYRLHPLPISDLKAIQVAREITLHTAP 923
Qy 1285 PDWTDYNDFELQMLPRDALTVFNPNYNNKSEGRIVYGLRPGRSYQFNKTVSGDSWKI-- 1342
Db 924 AG--EYTDLFQYLSAEEAPQLQNVTKNTEITLQGLRPYHNVYTFVWVRSIGIQGTF 981
Qy 1343 -----YSKPIFGSVRT---KPKIQNLHCRPQNSTAIACSW-IPP---DSDFDGY 1385
Db 982 ADVSVTLMRSSAPISASYQTLTAPPQGVVDYFQPSDVQPGEVTFEWSLEPAEQHGPIDYF 1041
Qy 1386 SIECKMDQVRFPSRKLKESLNTLMVPHKRYLVSIKVQSA-GWTSSEVVEDSTITM 1444
Db 1042 RITQONADDAADVSEYFPVNVATQKIDGLVPGNHVIFRQAKGALGVGAEREHIQWPI 1101
Qy 1445 IDRPPPPPHIRVNEKVDLISKSSINFVNCSEFSDTNGAVKYTVVVRADGSDDELKPE 1504
Db 1102 L---APPVPESVTPLEVSRTSSITIEISFRQGYFSAHGMVRSYTIIIAEDVGK----- 1152
Qy 1505 QQHPLPSVLYVRHNASTRYQVNTYFASKCAE--NP-----NSNSKSFNIKGAEMESLG-G 1557
Db 1153 ----IASGLEMPSQDQVAY-TWMLPYQAIPEYNPFLTNSGRKS-----SLEAEHLITGTA 1204
Qy 1558 KRDTQOKFCQGLPKPHTAVRISRAFTQLEDELFKFTKPLYSDDTFESLPITHESEPLF 1617
Db 1205 NCDKHQAGYCNGLRAGTYYRIKIRAF--DED-----KFTDVTYSSPITTERSDTV 1254
Qy 1618 GAIEGVSAGLFLIGLVAVVALLCROKVSHGRERPSARLSIRDRPLSVHLNLGQGNR 1677
Db 1255 IVAATVSA-VLLVAVLVVV---YQHRQ--LIRASKLARMQDELAAL-----PSGYI 1303
Qy 1678 KTSCKPIKINQPEGHFMKLOADSNYLLSKRYBELKDVGRNQSCDIALLPENRGNRYNNIL 1737
Db 1304 TPNRPVHVQDPSEHYRIMSADSPRFSEEFELKHVGRDQACSPANLPNRPKNRFTNII 1363
Qy 1738 PYDATRVKLSNVDDPCSDYINASYICGNPFREYIVTQGLPCTKDDFWKWWQEVNHN 1797
Db 1364 PYDSHFRKLPQVDDDDGSDYINANMPGHNSPREFIVTQGFHSTRFEFWRMCWESRA 1423
Qy 1798 IVMTVQCEKRVKCDHYWPAQDQSLYVDLILQMLSESVLPEWTIREFKI---CGBEQL 1854
Db 1424 IVMLTRCFEKREKCDQYWPVDRVAMPFGDIKVQIIDTHYDHSISEFVSRNC----- 1478
Qy 1855 DAHLIRHPHYTVWPDHGVPTTQSLIQVTRVDYINRSPAGTVPVHCSAGVGRGTGF 1914
Db 1479 -ESRIMRHFHTTWFDFGVPPEPLSLRVFVRAFRDVTGD--MRPIIVHCSAGVGRSGTF 1535
Qy 1915 IALDRILQOLDKSDVDIYGAVHDLRLHRVHMVQTEQYVYLHQCVRDVLARK 1968
Db 1536 IALDRILQHIHKSVDYDIFGIVFAMRKERVFMVQTEQYVCIHQCLLAVLEGKE 1589
```

RESULT 7  
138670

protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type J precursor - human  
N;Alternate names: density enhanced phosphatase-1 (DEP-1); protein-tyrosine phosphatase  
C;Species: Homo sapiens (man)  
C;Date: 01-Mar-1996 #sequence\_revision 08-Mar-1996 #text\_change 09-Jul-2004  
C;Accession: 138670; 152599  
R;Osman, A.; Yang, Q.; Tonks, N.K.  
Proc. Natl. Acad. Sci. U.S.A. 91, 9680-9684, 1994

A;Title: Expression of DEP-1, a receptor-like protein-tyrosine-phosphatase, is enhanced.  
A;Reference number: 138670; MUID:95024024; PMID:7937872  
A;Accession: 138670  
A;Molecule type: mRNA  
A;Residues: 1-1337 <RES>  
A;Cross-references: UNIPROT:Q12913; UNIPARC:UPI000017383B; EMBL:U10886; NID:G558754; PID  
A;Experimental source: HeLa cells  
R;Honda, H.; Inazawa, J.; Nishida, J.; Yazaki, Y.; Hirai, H.  
Blood 84, 4186-4194, 1994  
A;Title: Molecular cloning, characterization, and chromosomal localization of a novel pr  
A;Reference number: 152599; MUID:95086212; PMID:7994032  
A;Accession: 152599  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-216, 'LFGVKKAA', 225-260, 'G', 262-285, 'GTEGLDASNTERRA', 302, 'S', 304, 'TAPVHDE  
A;Cross-references: UNIPARC:UPI000016AF3B; GB:D37781; NID:IG633072; PID:BA07035.1; PID:  
A;Comment: Enhanced expression of this protein with increasing cell density suggests a  
C;Genetics:  
A;Gene: GDB:PTPRJ  
A;Cross-references: GDB:385040; OMIM:600925  
A;Map position: 19q13.4-19q13.4  
C;Function:  
A;Description: catalyzes hydrolysis of peptidyl-phosphotyrosine to peptidyl-tyrosine and  
C;Superfamily: protein-tyrosine-phosphatase, receptor type J; fibronectin type III repea  
C;Keywords: glycoprotein; phosphoprotein; phosphoric monoester hydrolase; receptor; tran  
F;1-35/Domain: signal sequence #status predicted <SIG>  
F;36-1337/Product: protein-tyrosine-phosphatase, receptor type J #status predicted <MAT>  
F;118-197/Domain: fibronectin type III repeat homology <3FNA>  
F;206-283/Domain: fibronectin type III repeat homology <3FNB>  
F;284-356/Domain: fibronectin type III repeat homology #status atypical <3FNC>  
F;355-445/Domain: fibronectin type III repeat homology <3FND>  
F;453-530/Domain: fibronectin type III repeat homology <3FNE>  
F;539-617/Domain: fibronectin type III repeat homology <3FNF>  
F;720-804/Domain: fibronectin type III repeat homology <3FNG>  
F;972-988/Domain: transmembrane #status predicted <TM>  
F;1065-1287/Domain: protein-tyrosine-phosphatase homology <PTP>  
F;72,82,93,104,142,172,192,231,258,278,342,351,376,391,396,413,431,501,525,536,582,603,6  
F;1239/Active site: Cys (phosphocysteine intermediate) #status predicted  
F;1245/Binding site: substrate phosphate (Arg) #status predicted

Query Match 11.9%; Score 1249; DB 1; Length 1337;  
Best Local Similarity 28.9%; Pred. No. 3e-61;  
Matches 417; Conservative 215; Mismatches 547; Indels 264; Gaps 61;  
Qy 641 TVPSSVSGVTNNSSGRNDYLSVSLVAPGDVDNYEVTLSDHGKVVQSLVIKSVRECSFS 700  
Db 50 TVATGNGITQISSTAESFKQGTGTF---QVETNTSEDG---ESSGANDSLR----- 97  
Qy 701 SLTPGRLTYVTITRSKYNHNSFSQF---RTVPDKV---QGVSVNSARSVDYLRVSWH 754  
Db 98 -----TPEQSGNCTDGASQKTPSPSTGSPVFDIKAVSISPT---NVILTWS 141  
Qy 755 ATGDFDHYEVTIKNNFIQTKSIKSENE-----CVFQVLPGRLYSVTVTTKS 804  
Db 142 NDTAASEYKVVKH-----XNENKTIIVHQPCNITGLPATSYSFISIPGI 190  
Qy 805 GOYEANEQ-NGRGTI-----PEPVKDLTLNRSTEDLHVTWSGANG-----DVDQY 849  
Db 191 G----NETWGDPRVIVTEPIPVSDLRVAHGCEGCSLSWSNGNGTASCRLVLESIGSH 246  
Qy 850 EIQLLFNDKMKVFPFHLVNTATYRFTSLTPGQYKILVLT-----SGD---VQOSAFIE 902  
Db 247 E--ELTQDSRL-----QVNI5DLKPGVQYINPYLLQSNKTKGDPLOAKVAM-- 291  
Qy 903 GFTVP5AVKNIHISPNGATDSLTVNWTGPGDGDVDSYTVSAFRHSQKQSDOTIPKHVEHT 962  
Db 292 ---MPALQRE---AGQAPPLCM-MSPFVGPVDPSS-----GQOSRDTVEL----- 331  
Qy 963 FHRLEAGEQYQIMTAS--VSGSLKNQINNVGRTPVPSVQGVVIADNAYSSYSLIVSWQKAA 1020  
Db 332 LVGLEPCTRYNATVYSQAANGTEGQPOAIEFRINAIOVFDVTAVN-ISATSLTLIKVSD 390  
Qy 1021 GVAB---RYDILLTENGILLRNTSEPTATTKQHKFEDLTGPKK-----YKILTVSGGLP 1073

```
Db 391 NESSNNTYKIHVAGETDSSNLNVSEPR-----VIPGRSSTFYNITVCPVLGDIE 442
Qy 1074 SKEAQTEGRTVPAAVTDLRITENSTRHLSFWTASEGELSWYNIFLYNPDGNLQERQVD 1133
Db 443 GTPGFLOVHTPPVPVDFRVTVVSTTEIGLAWSHDAESFQMH-----TQEGAGNSRVEI- 498
Qy 1134 PLVQSFQNLLOGRMYKMWIVTH--SGELSNESFIFGRTPVPSVSHLGRSNRTTDSLW 1191
Db 499 TTNSIIGLFPQTKYCFEIVKPGNGTEGASRTVCNRTVPSAVFDIHVVYVTTT-MW 557
Qy 1192 FNNWSPASGDFDF-YELILYNPNKGNKWKDOLTEWRFOGLVPRKYVLWV-----VTHSG 1247
Db 558 LDWKSPPGASEVYVHLVIESKHGSHNTSTYDKAIT--LQGLIPCTLYNITISPEVDHW 614
Qy 1248 DLSNKVTAESRTASPSLSMFADIANSTLAIWKGPDPDWDYNDYDFLQWLPRDALTVFN 1307
Db 615 GDNSTAQYTR-----PSNVNIDVSTNTTAATL-----SWQFDDASPTY-----SYCLLIE 662
Qy 1308 PYNRKSEGRIVG-----LRPGRSYQFNVKTVSGDSWKTYSKPIFGSVRTKPDK 1357
Db 663 KAGNSNATQVVDIGITDATVTELI FGSSYVVEIFAQVGDGKSL-EPGRKSECTDPAS 721
Qy 1358 IQNLHCR-PONSTAIACSWIPDSDFDGYSIE-----CRKMDTQEVFSFRKLEKEKS 1408
Db 722 MASPDCEVWPKEPALVLKWTCPGANAGFELEVSSGAWNNATHLESCESENGTEYRTEVT 781
Qy 1409 LLNIMLVLPHKRYLVLSKVQSGAGTSEVVEDSTITMTDRPPP--PHIRVNEKDVLSK 1466
Db 782 YLNF-----STSNISITVSCGMAAPTRCTGTGTPPPDGPSPNI-----TSVSH 830
Qy 1467 SSINFTVNCWFDNGAVKYFTVVREADGSDDELKPEQOHPPLPSYLEYRHNASIRVYQT 1526
Db 831 NSVK--VKFSGFEASHGPIKAYAVILTG-----EAGPSADVLYKYTYDFDKCASD 880
Qy 1527 NY--FASKCAENPNNSKS-----FNIKLGAPMESLGKRDPTQOKFCDGPLKPHATYRIS 1580
Db 881 TYVTYLIRTEBKGRSLSLEVLKYEIDVGNESITLG-----YNGKLEPLGSYRAC 931
Qy 1581 IRAFTQ-----LFD--EDLKEFTKPLYSDFTFPSLPITTSSEP--LFGATEGVSAGL 1627
Db 932 VAGFTNITFHPQNKGLIDGAEVYVSFSR--YSDA-VSLP-----QDPGVICGAVFGCIFGA 984
Qy 1628 FLIGMLVAVVALLICROKVSNGRERPSARLSIRRDPLSVHLNLGQGNKRTSCPIKINQ 1687
Db 985 LVI---VTVGGFIFWRKK--RKDAKNEVSFSQIKP-----KSKLIRVEN 1025
Qy 1688 FEGHMKLOADSNYLLSKYEELKOVGRNQSCDIALLPENRGKRNRYNNILPYDATRYKLS 1747
Db 1026 FEAYFKKQADNSCGFAEEYEDLVGISQPKYAABLAENRKNRKNRYNNVLPYDISRVKLS 1085
Qy 1748 NVDDPCSDVINASYIIPGNFRREYIYVQGLPGTKDDFWKWVWQNVHNVMTQCEK 1807
Db 1086 -VQTHSTDDYINANYPGYHSKKDFIATQGLPNTLKDFWRWVWKNVYAIIMLTCKVEQ 1144
Qy 1808 GRVKCDHWPADQDLSYGLLILQMLSESVLPWTIREFKICEEQLDAHLRTHPHYTV 1867
Db 1145 GRTKCEYNSKO-AQYGGITVAMTSEI VLPEWTRIDFTVKIQTSESHPL-RQHFSTS 1202
Qy 1868 WPDHGVPTTQSLIQFVRTYRDYINRSPGAGTPVHCSAGVGTGTFIADRLILQQLDSK 1927
Db 1203 WPDHGVPTTDLINFRYLVRDYMKQSPSPILVHCSAGVGTGTFIADRLIYQIENE 1262
Qy 1928 DSVDIYCAVNDLRLHRVHMVQTECOYVYLHQCDVDVLARKLS-----EQENLFPFIYEN 1983
Db 1263 NTVDVYGIYDLRMHRLPMVQTEQYVFLNQCVLDIVRSQKDSKVLDIYQNTTAMTIYEN 1322
Qy 1984 VNP 1986
Db 1323 LAP 1325
```

```
S68700
HPTP beta-like tyrosine phosphatase precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S68700
R:Kuramochi, S.; Mateuda, S.; Mateuda, Y.; Saitoh, T.; Ohaugi, M.; Yamamoto, T.
FEBS Lett. 378, 7-14, 1996
A:Title: Molecular cloning and characterization of BYP, a murine receptor-type tyrosine kinase
A:Reference number: S68700; MUID:96140699; PMID:8549806
A:Accession: S68700
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1238 <KUR>
A:Cross-references: UNIPROT:Q64455; UNIPARC:UPI00000013C2; GB:D45212; NID:g1208432; PIDN
C:Genetics:
A:Map position: 2E1-2
C:Superfamily: protein-tyrosine-phosphatase, receptor type J; fibronectin type III repeat
C:Keywords: phosphoprotein
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-1238/Product: HPTP beta-like tyrosine phosphatase #status predicted <MAT>
F:267-347/Domain: fibronectin type III repeat homology <3FR>
F:966-1188/Domain: protein-tyrosine-phosphatase homology <PTP>
F:1140/Active site: Cys (phosphocysteine intermediate) #status predicted
F:1146/Binding site: substrate phosphate (Arg) #status predicted

Query Match 11.9%; Score 1243.5; DB 2; Length 1238;
Best Local Similarity 29.4%; Pred. No. 5.3e-61;
Matches 394; Conservative 194; Mismatches 481; Indels 269; Gaps 53;

Qy 750 VSWHATGDPDHVEVTKKNNFIQTKSPKSENECVFVOLVPGRLYS--VTVTTKSGQ 806
Db 57 LTWKH--NDSCASECRIENKMSNLTPPV-KNOTSCNITGLSPGTSYTFISIIVTT--- 109
Qy 807 YEANEQNGRTI---PEPVKDLTLNRSTEDLHVTHSGANGDVDDYEQILFLDNMKVFP 863
Db 110 ---NETLN-KTITTEPWPVSDLHVTSVGTQABLTVSNANGTA---SYRMLIEEL--- 157
Qy 864 FHLVNTATEYRFTSLRPGROYKILVLTISGDVQOSAFIEGTFTVPSA--VKNIHIS--- 916
Db 158 ---TTHSSVNIISGLKPGTN-NTFAPESNET-QADFAVAEEVPDANGTKRIPTNLSQL 211
Qy 917 -----PNGATDSLTVNWTGPGGDVDSYTVSAFRHSQKVDSTQIPKHVFHEHTFHL 966
Db 212 HKNSLVSDPPSGDPSLT-----EILLTDL 237
Qy 967 EAGEQVQIMTAS--VSGSLKNQINNVGRTVPASVQGVADNAYSS$LYISWQK----- 1018
Db 238 KPDTQYNATIIYQOANGTEGQPRNKVFKTNSTQVSDVRAMN-ISASSMTLITWKSNDYDGR 296
Qy 1019 -----AAGVAERYDILLTENGILLRNTSEPTATTKQHKPEDLTPGKKYKIQLTVS 1069
Db 297 TSIVYKIHVAGGTHSVNQTVNKTETAILGLSSSTLYNITVHPFLGQTGTPGLQVY--- 353
Qy 1070 GGLFSKEAQTEGRTVPAAVTDLRITENSTRHLSFRWTASEGELSWYNIFLYNPDGNLQER 1129
Db 354 -----TSPQVSDPRVTNVSTRAIGLAWRSNDSKS--PEIFIKQDGGEKHN 398
Qy 1130 AQVDPVLVQSPFQNLLOGRMYKMWIVTH--SGELSNESFIFGRTPVPSVSHLGRSNRTT 1187
Db 399 ASTGN--QSYMVSDLPKPGTSYHFEIIPRPGDGTGELSSSTVNGSTGDSAVTDIIRVNIIST 456
Qy 1188 DSL--WFWNSPASGDFDFYELILYNPNGT--KKNWKKDOLTEWRFOGLVPRKYVLWV-- 1242
Db 457 EMGLEWQNTDASG--YTHLVLESKSGSIIRTNSSQKWIT---VGSLTGTLNVTIIPP 511
Qy 1243 -VTHSGDLGNKVTAESRTASPSLSMFADII--ANTSLAITWKGPDPDWDYNDYDFLQWL 1299
Db 512 EVDIQIGISNITQYTR-----PSSVSHIEVNTTTTAAIRWKN----- 550
Qy 1300 RDALTVPNPY-----NNRKSEGRIVVGLRPRGRSYQFNVKTVSGDSWKTY 1344
Db 551 EDAASASYAVSVLILKTDGSGNVTNNTKDPSPILIPELIPGVSYTVKILTVQVGDG-TTSL 609
```

Query Match	11.4%	Score 1198	DB 1	Length 1711	
Best Local Similarity	25.3%	Pred. No. 3.1e-58			
Matches	411	Conservative 256	Mismatches 657	Indels 302	Gaps 53
Qy	477	HANETSLSIMWQTPVAEWKYYIISLADRDLLLIHKSLSKDAKE-----FTFTDLVP	527		
Db	46	HGKSTSLFUSWAAELGGDFYALUSRVD-----SSGSPGQQLQAHTNMSGFEFHLGP	100		
Qy	528	GRXYMATVTSISGDLKN-SSSVKGRTPVAQVTDLHVANQGMTSSLFTNWTQAQDVFYQ	586		
Db	101	GSRYQLKLVLRPCWQNVITLTARTAPTVRGQLHSGASPARLEASWSDAPGDQDSYQ	160		
Qy	587	VLLIH--ENVVIKVESISSTSYSHLSKSGSLYSVVVTVTSGGISRSRVVVEGRTPVS	644		
Db	161	LLLYHLSESQTLACNVSVSPDLSYSGFDLLPGTYQVLEVITWAGSLHAKTSILQW-TEPV	219		
Qy	645	SVSGVTVNNSGRNDYLSVSWLVAPGDVNDYEVTLSDHGKVVQVSLVIKASVRECSFSLTP	704		
Db	220	PPDHLALRALGTS--LQAFWNSSEGATSFHMLTDLGGTNTTAVIRQGVSTHTFHLSP	278		
Qy	705	GRLYVTVTITRSKYYENHSFQERTVPDKVQGVSVNSARSDY-----LRVSW	752		
Db	279	GTPEHKICASAG-----PHQIWGFSATEWTVPSPSDLVLTPLRNLWASW	325		
Qy	753	VHATGDFHYEVTIRKNNPFIOTKSIKSENECFVQLVPGRLYSVTVTTKSGQYANEQ	812		
Db	326	KAGLGARDGY--VLKSLGPMESTLGPBECNAVFPGPLPGHYTLQLKVLGAGPYDAWVE	383		
Qy	813	GN-----GRTIPBPVKDLTLNRSTEDLHWV--SGANGDVQVIEIQLLFNDMKVPPFHL	866		
Db	384	GSTWLAESAALPREVPGARL-----WLDGLEASKPQRRALLYSD---DAFGSL	429		
Qy	867	VNT-----ATEYFTSLTPGRQYKILVLTISGVDVQQSARFEGTVPSAVKNIHSPNGAT	921		
Db	430	GNLSVPSGATHVIFCGLVPGAHVRDIASGTGISQ--ISGTSPLPPQSLEVISRSP	487		
Qy	922	DSLTVNMTGGGDVDYTVSAPRH-SQKV--DSQTIPIKHVFERTFHRLEAGEQYQMIA	978		
Db	488	SDLTIAWGPAPQLEGYKVTWHQDGSORSPGDLVDLGPDTLSLTLSLVPGSSYTVS	547		
Qy	979	VSGLL-KNQINVGRTPASVQGVADNAYSYSLVSWQAKAQAERYDILLTENGIL	1037		
Db	548	WAGNLGSDSKIHSCTRPAPTNLSLGFAPHAALKASWYHPFGGRDAFHLRYLRPLT	607		
Qy	1038	LRNTSE-PATTKQHKEDTTPGKKYKIQILTVSSGGLFSKEAQTEGRTPAAVTDLRITEN	1096		
Db	608	LESEKVLPREAQNFSAWLTAGCEFOVQUSTLWGSERSSANATGWTTPPAPTLVNTSD	667		
Qy	1097	STRHLSFRWTASSEGLSWYNIFLYNPDGNLQBRQAVDPLVQSFSSFQNLGGRMYKMVIT	1156		
Db	668	APTQLQVMAHVPGGRSRQVTLYQ-ESTRTATSIMGPKEDGTSFLGLTPGKYKVEVIS	726		
Qy	1157	HSGELNESPIFGRTVPASVSHLRGSRNRTDLSLFWNWSPASGDFPIYELILYNPGTKK	1216		
Db	727	WAGPL-----YTAANVS-----AWT-----YPLI---PN-----	748		
Qy	1217	ENWKDXDLTEWRFGQLVPGRKYYLVWTVHSHGDLSNKVTAESRTAPSPPSLMSADTANTS	1276		
Db	749	-----ELWVSMOAGS-----AVVNLAWPSGP-LGQGACHAQLS	780		
Qy	1277	LA--ITWKGPPDWDYNDFELQWLRDALTVFNFPYNNRKSEGRIVYGLR---PGRSYQFN	1331		
Db	781	DAGHLSWEQ-----LKGQELFMRLDLTFGHTISMS	812		
Qy	1332	VKTVSGDSWTKYSKPIFGS-----VTKPKDKIQNLHCRPONSTAIACSWTPPDSPDGY	1386		
Db	813	VRCRAG-----PLQASTHLVVLVSVEPGPVEDVLCHPE-ATYALNMTMPAGDVDCVL	863		
Qy	1387	IEC-RKMDTQEVFSRKLKESLLNIMLVPHKRYLVLSIKVQSAGMT-SEVVEDSTITM	1444		
Db	864	VVVERLVPGGGTHFVQVNTSGDALLPLNIMPTTSYRLSLTVLGRNSRWRASVLCVSTS	923		







```
Qy 1181 --GSRNRTTDSLWFWNPASGDFDFYELILYNPNGTCKENWKDKDLTEWRFOGLVPGKY 1238
Db 470 PIGNDETDTKLTKESTPKNIDVPF-----AKNLVKACY 504
Qy 1239 --VLWVTHSGDLGNKVTAERTAPSPPLMSFADIANTS--LAITWKGPPDWTYNDFE 1294
Db 505 RVQLFTVTKTGIISETRHNETIRMSPPAVNVLSLSTRSSATLRIVFSTHDSISNCQ 564
Qy 1295 LQWLPRD-----ALTVPNPNKRKSEGRIVVGLRPGRSYQFNKVTSGD--- 1338
Db 565 MHIVRDMNGKSVFDMKQLTATFAPLLN-----LDGLSPFHXYTNTQIICGSGSE 617
Qy 1339 -----SWKTSKPIFGSVRTKPKIOHLCRPNSTAIACSWLP---PDSDFDGYIEC 1389
Db 618 TPQCPAATRTNRQLSFSTRQDKPAVQDLKVEPLNSVMLTWLPPALPNGLTHYAVNV 677
Qy 1390 RKM---DTQEVFERKLEKEKSLNIMW--LVPHKRYLVSIK--YQSAGMTSEVVEDSTIT 1443
Db 678 TKIGSDTRTIDVGVSNRSDHTVQVVIDELFGGHTYSFSVRAVTEAGFGENSPPVPTVS 737
Qy 1444 MIDRPPPPPHIRVNEKDVLSKSSINFTYN--CSWFSDTNGAVKYFTVVVREADGSDDEL 1501
Db 738 MPLMAPPVPTVAPMIMKE---SVGSHNMIVRFPTTFMFDNRNGEIKQFAIIVSETTADESI 794
Qy 1502 -----KPEQHQPLPSYL-----EYRHNASIRVYQTVNFASKCAENPSN 1540
Db 795 NRWIESDNGTYTWOQVQRFVDFWPSYVAKLQDIQVKQDVDSIPE----- 839
Qy 1541 SKSFIKILGAMESLGKGRDPTQOKFCGDLKPKHTAYRISIRAPTLQFDEDLKEFTKPLY 1600
Db 840 -----ELGEDETCLEVRAD-----RICNGPLRSASKVRVIRLFT-----SPTLF 879
Qy 1601 SDTFESLPITIESE-----PLFG-----AIEGVSAGLF----- 1629
Db 880 TDSPPSQMTTGSATPAIPLTVVAVLIVAFVIGVITFLFFWNRTKXKARLAAAFKNG 939
Qy 1630 -----IGMLVAVALLICROKVSGRHREPSARLSIRDRDPLSV----- 1667
Db 940 PSKEKESQWEALKMMAERA-ADCLAKLGLDATTPPSSSTSSNSTSTSTMTDCGSP 998
Qy 1668 HL---NLGQKNRKT-----SCPTIKNQ 1687
Db 999 HLGAPNAG--CHRRTRSLRERTGVEHRLERLSSGPVHRTPLYTVVTGANTKSRPVRIED 1056
Qy 1688 FEGHFMKLQADSNVLLSKYEELKQCRNQSCDIALLPENRKNRYNNILPYDATRKLS 1747
Db 1057 FADHVLMSADSDFRFSEEDMMRNVGVSVAASELPINRPNKRNFTNIPSDHSRKLS 1116
Qy 1748 NVDDDCSDVINASYIPGNFRREYIVTQGPLGTGDDFWKVMQNVHNIVMVTOCVEK 1807
Db 1117 NPNNIEGDDVINANYVFGFSRREFFIAAQGPLPTTRDFHFWQMTWEQOCPAIIALTCKVEK 1176
Qy 1808 GRVKCDHYPADQD-SLYYGDILQMLSESLPEWTIREPKI--CGBEQDLDAHLRIHFH 1864
Db 1177 GROKCHQWPDHENVPLVYGDIEVTIIVAEKEFEFVIRDIRELSKSPDG-RVTRFVRHWH 1235
Qy 1865 YTWPDHGVETTSQSLIQFVTRVDYINRSPGAGPTVWHCSAGVGRGTPTIALDRILQQL 1924
Db 1236 YMAWPDFGAPSHPNIGIIQFSRMFRHLPHSPHNAPTIVHCSAGVGRSGTPTISIDRLQSS 1295
Qy 1925 DSKDSVDIYGAVDLRLHRRVHMVQTEQYVYLHOCVRDVLRL 1965
Db 1296 SFGDPIDVFGTVCBMYRCMQVNEQQYIFIHYCILQVLQ 1336
```

RESULT 12  
A57064

protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type O precursor - human  
N;Alternate names: GLEP1; glomerular epithelial protein 1  
N;Contains: protein tyrosine phosphatase phi, cytosolic form; protein tyrosine phosphatase  
C;Species: Homo sapiens (man)  
C;Date: 03-Oct-1995 #sequence\_revision 08-Feb-1996 #text\_change 22-Jun-1999

C;Accession: A57064  
R;Wiggins, R.C.; Wiggins, J.E.; Goyal, M.; Wharram, B.L.; Thomas, P.E.  
Genomics 27, 174-181, 1995  
A;Title: Molecular cloning of cDNAs encoding human GLEP1, a membrane protein tyrosine p  
ne to human chromosome 12p12-pl3.  
A;Reference number: A57064; MUID:95394455; PMID:7665166  
A;Accession: A57064  
A;Status: not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-1188 <WIG>  
A;Cross-references: UNIPARC:UPI000004DC16; GB:U20489; NID:G895925; PIDN:AAA82892.1; PID:  
C;Genetics:  
A;Gene: GDB:PTPPO  
A;Cross-references: GDB:454477; OMIM:600579  
A;Map position: 12p13.3-12p13.1  
C;Superfamily: protein-tyrosine-phosphatase, receptor type O; fibronectin type III repeat  
F;1-29/Domain: signal sequence #status predicted <SIG>  
F;30-819/Domain: extracellular #status predicted <EXT>  
F;30-109/Domain: fibronectin type III repeat homology #status atypical <FN3A>  
F;116-202/Domain: fibronectin type III repeat homology #status atypical <FN3B>  
F;329-409/Domain: fibronectin type III repeat homology <FN3C>  
F;432-520/Domain: fibronectin type III repeat homology <FN3D>  
F;529-619/Domain: fibronectin type III repeat homology <FN3F>  
F;631-714/Domain: fibronectin type III repeat homology <FN3G>  
F;722-804/Domain: fibronectin type III repeat homology <FN3H>  
F;812-1188/Domain: product: protein tyrosine phosphatase phi, long form #status predicted <PHIL>  
F;812-876,877-1188/Domain: product: protein tyrosine phosphatase phi, short form #status predicted  
F;820-844/Domain: transmembrane #status predicted <TMN>  
F;845-1188/Domain: intracellular #status predicted <INT>  
F;902-1188/Domain: product: protein tyrosine phosphatase phi, cytosolic form #status predicted <I  
F;934-1156/Domain: protein-tyrosine-phosphatase homology <PPT1>  
F;75,154,189,201,227,278,287,323,324,370,461,490,700,712,733,790/Binding site: carbohydra  
F;1108/Active site: Cys (phosphocysteine intermediate) #status predicted  
F;1114/Binding site: substrate phosphate (Arg) #status predicted

Query Match 8.7%; Score 913.5; DB 1; Length 1188;  
Best Local Similarity 24.0%; Pred. No. 1.2e-42;  
Matches 367; Conservative 199; Mismatches 483; Indels 481; Gaps 63;

Qy 571 LFNWTOAGDVEFYOVLLIHENVVKNES-----ISSETGRY----- 608  
Db 23 LFNKAT-----AFHVTQDDNNIVVLSLEASDVISPASVVVVKITGESKNYFFEFENFS 76  
Qy 609 -----SPHLSKSGSISYVVVTVTSGGI---SSRQVVVBGRTPVSSVSGTVNNSG 655  
Db 77 TLPPPVIFKASHY---GLYIIITLVVNGNVVTKFSRSTTV---LTKPLPVTSVSYDYK 130  
Qy 656 RNDYLSVSWLVPDGDVDNY--EVTLSH-DGKVYQSLVI-----AKSVRECSFSLTPGRL 707  
Db 131 PSPTGVLPFIHYPEKYNVFTRVNISIYWGKDPRTMLYKDFPKGKTIV---FNHMLPGMC 186  
Qy 708 YT-----VTITRSGKYNHSFGSQRTPDKVQGVSVNSARSYDLRV--- 750  
Db 187 YSNITPOLVSEATFNKSTLVEYSG--VSHEPKQHRTAPYPPQNISV-----RIVNL 235  
Qy 751 ---SWHATGDPDHYEVTIKKNKFNFTQTSIKPSENECVFVLVP-----GRLYSVTVTK 803  
Db 236 NKNWEEQSGNFP-----BESFMRSDTIGKEKLFHTEETPEIPSGNISGWPDFN 287  
Qy 804 SGQYEANEQ---GNRTIPEP-----VKDLITRNRTEDLHVTWGSANGVDQYEIQLLF 855  
Db 288 SSDYETTSQPYWDSASAPESDEFEVSVLPMEYENNSTLSEKTSKSGSFSPFPQMIL 347  
Qy 856 NDMKVPPP-----FHLVNTATEYRFTSL-----TPGRQYKILVLTISGDV 895  
Db 348 TWLPPKPPPTAFDGFH-IHIERENFTEYLMVDDEAHEFVAELKEPGK-YKLSVTTFSS--- 402  
Qy 896 QQSAPTEGTVTSVAVK-NIHSIPNGATDSLTVNWTGGGDVDSYTVSAFRHSQVDSQTI 954  
Db 403 -SSGSCETKRSQSAKLSYIPSPG-----EW-----IEBELTEK----- 435  
Qy 955 PKHVFHTHRLRAGEQYQIMIASVSGSLKNQINNVVGRTPVSPASVQGVQVIADNAYSYSLI 1014



Db 436 PHSVSVHLSTTA-----LMSWTSQENYSTI-----VSVSLTC 472  
Qy 1015 SWQAAAGVAERYDILLTENGILLRNTSEPAATKQHPEDITPKQYKIQILTVSGGLFS 1074  
Db 473 QKQKE---SORLEQYCTQ-----VNSKPI-----IENLVPAQYQVVIYLRKGLIG 518  
Qy 1075 KEAQ-TEGRTPAAVTLRITENSTRHLSFWTASE-CELSWY--NIFLYNPDGNLQERA 1130  
Db 519 PPSDPVTPAIYPTGIKMLYPLGPTAVLVSWTRPYLGVFRKYVYEMFYFNPATMTSEWT 578  
Qy 1131 QVDPVLQVFSFQNLQGRMYKXVIVTHSGELSNESFIFGRVTPASVSHLRGSRNNTDLSL 1190  
Db 579 TYEIAATVS-----LTASVRAN-----LLPA----- 601  
Qy 1191 WFNWSPASGDFDYELILYNPNGTCKENWKDKJLTFWRFGQLVPGRYKVLWVTHSGDLS 1250  
Db 602 WY-----YNRVPT-MVTWGDPELS-----CC 621  
Qy 1251 NKVTASRTAPSPSLMSFADIANSTLAIWKGPDPDWDYDNFB-LQWLPDALTVFNPY 1309  
Db 622 DSSSTISFIPAPVAPEITS-VEYFNSLLYISWYGDYDLDLHSHRMLHWM----- 669  
Qy 1310 NNRKSEGRIVYGLRGRSYQFNWKTVDGSKVSKPIFGSVRT-----KPKIQNLHCR 1364  
Db 670 -----VABGKKIKKSVTRVMTAILSLPDPDIYNL--- 701  
Qy 1365 PNSTAIACSWIPDPDGYGSIKCKMDTOEVEFSRKLKESKLLNIMMLVPHKRYLVS 1424  
Db 702 ---SVTACT-----ERGSNTSMRLVKL----- 721  
Qy 1425 IKVQAGMTSEVEDSTIMIDRPPPPHTRVNEKOVILSKSINFVNCWSFSDTNGA 1484  
Db 722 -----EPAPPKSLFVANKTQ-----SVTLWYEE--GV 748  
Qy 1485 VKYFTVVVREAGDDELKPEQHPPLPSYLEYRHNASIRVYQTNYPASKCAENPNSNSKSP 1544  
Db 749 ADFFEVPCQVGSQKTKQBPVAVS-----HVVTI----- 780  
Qy 1545 NIKLGAEMESLGGKRDPTQKQCDGPKLPHTAYRISIRAFATQDFEOLKEFTKPLYSDF 1604  
Db 781 -----SSLLPATAYNCVTSFSH-----DSPSV-PTF 806  
Qy 1605 FSL-PITTESPLFCAIGVS-AGLFLIGMLVAVALLICRQK-VSHGRERPSARL----- 1657  
Db 807 IAVSTMVTEMNPNVWVIVSLAILSTLITGLL--LVTLIILKKHLQMAECGAGTFVNF 864  
Qy 1658 SIRDRLPLSVHLNCGQGNRTSCPIKINQFEGHFMKLOADSNYLLSKEYEELKDVGRNQ 1717  
Db 865 SLERDGLPYNWSKNGLKKRKLTPVQLDDFDVAYIKOMAKDSYKFSIQFBEELKILGDI 924  
Qy 1718 SCDIALLPENRGNKRNILPYDATRVKLSNVDDPCSDYINASVYIPGNFRREYIVTQ 1777  
Db 925 PHFADLPLNCKRYTNILPYDSRVLVSMVEEGADYINANYIPGYNFPOBYIATQ 984  
Qy 1778 PLPGTKDDFWKQWVEQNVHNIWMTVQCVEKGRVKCDHYWAPADQDSLYYGDILQWLSGV 1837  
Db 985 PLPETENDFWKWLQKQSQIIVMLTQCNEKRVKCDHYWPPTEPIAYGDTIVEMISEE 1044  
Qy 1838 LPWTIREFKICGEOLDLHRLIRHFHYTVWPDHGV--ETQSLIQPRTVRYINRSP 1895  
Db 1045 QDDWACRHFRI---NYADEMDQVMHFNTAYTPDGHVPTANAAESILQFVHMVROQATKS- 1100  
Qy 1896 GAGPTVHCSAGVGRGTGFIALDRILQOLDSKSDYIYGAVHDLRLHRVHMVQTECOVY 1955  
Db 1101 -KGPWIIHCSAGVGRGTGFIALDRLLQHIRDHFVDILGLVSEMRYSRMSVWQTEQYIF 1159  
Qy 1956 LHQCVRDVLARKLRSEQENLFFIYENVN 1985  
Db 1160 IHQCQLMMWKKK---QQFCISDVIYENV 1186

RESULT 13

A53661

protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type O precursor - rabbit  
N:Alternate names: GUEPPI; glomerular epithelial protein 1  
C:Species: protein tyrosine phosphatase phi, cytosolic form; protein tyrosine phosphatase  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 07-Oct-1994 #sequence\_revision 08-Feb-1996 #text\_change 09-Jul-2004  
C:Accession: A53661

R:Thomas, P.E.; Wharram, B.L.; Goyal, M.; Wiggins, J.E.; Holzman, L.B.; Wiggins, R.C.  
J. Biol. Chem. 269, 19953-19962, 1994  
A:Title: GUEPPI, a renal glomerular epithelial cell (podocyte) membrane protein-tyrosine  
A:Reference number: A53661; MUID:94327545; PMID:7519601

A:Accession: A53661

A:Molecule type: mRNA

A:Residues: 1-1187 &lt;THO&gt;

A:Cross-references: UNIPROT:Q28613; UNIPARC:UPI0000086BEA; GB:U09490; NID:9529411; PIDN:  
A:Note: authors translated the codon GGC for residue 1101 as Gln  
C:Superfamily: protein-tyrosine-phosphatase, receptor type O; fibronectin type III repea  
C:Keywords: glycoprotein; kidney; phosphoprotein; phosphoric monoester hydrolase; recept  
F:1-29/Domain: signal sequence #status predicted <SIG>  
F:30-818/Domain: extracellular #status predicted <EXT>  
F:30-115/Domain: fibronectin type III repeat homology #status atypical <FN3A>  
F:116-209/Domain: fibronectin type III repeat homology #status atypical <FN3B>  
F:328-415/Domain: fibronectin type III repeat homology <FN3D>  
F:431-519/Domain: fibronectin type III repeat homology <FN3E>  
F:528-625/Domain: fibronectin type III repeat homology <FN3F>  
F:630-713/Domain: fibronectin type III repeat homology <FN3G>  
F:721-810/Domain: fibronectin type III repeat homology <FN3H>  
F:811-1187/Domain: protein tyrosine phosphatase phi, long form #status predicted <PHIL>  
F:811-875/Domain: Product: protein tyrosine phosphatase phi, short form #status predict  
F:819-843/Domain: transmembrane #status predicted <TMN>  
F:844-1187/Domain: intracellular #status predicted <INT>  
F:901-1187/Domain: protein tyrosine phosphatase phi, cytosolic form #status predicted <  
F:933-1155/Domain: protein-tyrosine-phosphatase homology <PTPI>  
F:75,154,189,201,227,277,286,323,369,460,489,699,711,732,789/Binding site: carbohydra  
F:1107/Active site: Cys (phosphocysteine intermediate) #status predicted  
F:1113/Binding site: substrate phosphate (Arg) #status predicted

Query Match 8.7%; Score 908.5; DB 1; Length 1187;

Best Local Similarity 25.0%; Pred. No. 2.4e-42;

Matches 366; Conservative 202; Mismatches 474; Indels 423; Gaps 67;

Qy 655 GRNDYLSVSL-VAPGDVDNVEVTLSDHGVQSLVIKSVRECSFSLTGLRLYTVTIT 713  
Db 10 GRRRLLELLMLFLVLLKTAFAHFVTRDNDNLSIVSL-----EAS-DVISASVYVVKIT 61  
Qy 714 TRSGKYENHSFSQER---TVPDKVQGVSVNSARSADYLRVSVWVHATGDFDHYEVIKKN 770  
Db 62 ---GESKNYFFEFENSTLPPPV-----IPKANY-----HGL---YYIITLVVN 101  
Qy 771 NFIQTKSIPKSENECVFQVLVPGRLYVTVTKSGQVEANEQNGRTRIPKPKDILNRR 830  
Db 102 GNVVTK-----PSR--SITVLTK-----PLPVTSVSIYDY 129  
Qy 831 STE-----DLHVT-----WSGANGDV---DQYEIQLLENDMKVFP--- 862  
Db 130 KPSPETGVLEIHYEYKYNVTRVNIWYEGKAPRTMLYKDFPKGKTVFNHW--LPGICY 187  
Qy 863 ---PFHLVNTAT-----EY-----RFTSLTFGRQYKILVLT---SGDV 895  
Db 188 SNITFQLVSEATFNKSTLVEYSGVSHPEKQHTAPYPRNISRIVNLKNWEEQSGSF 247  
Qy 896 QQSFAFISGFTVPSAVKNIHI-----SPNGATDSLTVNWTGCGVDVSTVSAFRHSQ 947  
Db 248 PEESFMRS---PETIEKDRIFHFTETPEPSG-----NISSGWDPFNS----- 287  
Qy 948 KVDSTQTPKHVFHTFHRLEAGEQYQIMIASVSGSLKKNQNVGRTVPASVQGVADNAY 1007  
Db 288 -SDYETTSQPYWMDASATPESEDEFVSLPMEYNNNTLSEAEKPTAPP-----SF 339  
Qy 1008 SSVSLIVSW--QKAAGVAERYDILLTENGILLRNTSEPAATKQ--HKF--EDITPGKKY 1061  
Db 340 FPVQMILSWLPPKPTAFDGFHIERE-----ENFTEYSTVDEAEHFAELKEPG-KY 393

```
Qy 1062 KIQILT--VSGGLFSKEAQT-----EGRTVPAAVTDLRITENSTRHLISFR 1104
Db 394 KLSVTFPSAGSCETRESQSAKLSFYISPTGEMIEELTERPQHVSFVHLSSTALMS-- 451
Qy 1105 WTASEGELSWNFIYLNPDG-NLQERAQVDPLVQSFSPQ-----NLQGRMYKRWI 1154
Db 452 WTSSQEN---YNSTIVSVVSTCQKESQRLKQYCTQVNSSKRIIENLVFGAQYQVM 508
Qy 1155 VTHSGEL---SNESIFIGRTVPASVSHLGRNSRNTTDSLWENMS-PASGDFDFV--ELL 1208
Db 509 YLRKGLPIGPSPDPVTF-AIVPTGKIDMLVPLGPT-AVLVSWTRPYLGVRKYVEMFY 566
Qy 1209 YNPNGTKKENW-----KDKDLT-EMRFQGLVPRKY---VLMVVTHSGDLS--NKVTA 1255
Db 567 FNP-ATWSEWTTYEIAATVSLTASVRIANLLPANYNFRVTWVTCDEPSCDSSTI 625
Qy 1256 ESRTAPSPPLMSFADIANSLAITWKGPPDWTYNDPE-LQWLPDRLTVFNPNRKS 1314
Db 626 SFITAPVAPEITS-VEYFNLSLYISWTYGDGDTTDLSHSRMLHM----- 668
Qy 1315 EGRIVYGLRGRSYQFNVTYSGDSWKTYSKPIFGSVRT-----KPKIQNLHCRPNST 1369
Db 669 -----VVAEGKKIKKSVTRNVMTAILSPGDIYNL-----S 701
Qy 1370 AIACSWIPPPDSDFDGYIECRKMDTQEVFESRKLKESKLINIMLVPHKRYLVSIKVQS 1429
Db 702 VTACT-----ERGSNTSMLRVKL----- 720
Qy 1430 AGMTSEVVEDSTIWMIDRPPPPPHIRVNEKDVLISSKSNFTVNCWFSSTNGAVKYFT 1489
Db 721 -----EPAPPKSLFVANKTQT-----SVTLWLWVEE--GVADFFE 752
Qy 1490 VVREAGDSDELKPEQHQHPLSYLEYRHNASIRVYQTYFASKCAENPNNSKSFNIKLG 1549
Db 753 VFCQVSGLETKLQEPVAVSS-----HVVTI----- 779
Qy 1550 AEMESLAGKRDPTQOKFCGDKPKPHTAYRISIRAFQTFDIEDLKFTKPLYSDTFSL-P 1608
Db 780 -----SSLLPATAYNCVTSFSH-----DSPSV-PTFIAVST 810
Qy 1609 ITTSEPLFGAIEGVS-AGFLIQLMVAVALLICROK-VSHGHERPSARL-----STRRD 1662
Db 811 MVTEMNPNVVISVLAISTLILGL--LVTLILRKKHLQMARECGAGTFVNFASLERD 868
Qy 1663 RPLSVHLNLQGNRKTSCTPIKNQFEGHFKMLQADSNYLLSKEYEELKDVGRNQSDIA 1722
Db 869 KGLPNYSKGLKRRKLTNPVQLDDFDAYIKDMAKDSYKFSQFEEKLKILGLDIPFAA 928
Qy 1723 LLPENRGNRYNLLPYDATRVKLSNVDDDDPCSDYINASYIPGNFRREYIVTQGPLGT 1782
Db 929 DLPLNRCKRYNLLPYDFSRVLLSNMEEGADYINANYIPGYNSPQEIYATOGPLPET 988
Qy 1783 KDDFWKMWQONVNIWVWTCQVEKRVKCDHYWPADODSLYIGDLILOMSSVLPEWT 1842
Db 989 RNDFWKMWLOQSKQMIWMLTQCNEKRVKCDHYWPFTEETIAYGDITVEMISEEQDWA 1048
Qy 1843 IREFKICGEQLDAHLRIJRHPHYTWPDHGV-ETTCQLIQFVTRVDRDYNRSPGAGPT 1900
Db 1049 HRHFERI---NYADEMQVMHFNFTAWPDHGVFTANAAESILQFVHMVRQOATKS--KGP 1103
Qy 1901 VVHCSAGVGTGTFTALDRILQQLDSDKSDVDIYGAVHDLRLHRVHMVQTCQVYVHLQCV 1960
Db 1104 IIHCSAGVGTGTFTALDRILQHLRDHEFDVILGLVSEMSRYMSVQTEBQIPIHQCV 1163
Qy 1961 RDVLRAKLRSQENPLFPYENNV 1985
Db 1164 QLMMWKKK---QQFCISDVIYENVS 1185
```

RESULT 14

S60613

protein-tyrosine-phosphatase (EC 3.1.3.48) U2 precursor - human

C;Species: Homo sapiens (man)

```
C;Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S60613
R;Seimiyu, H.; Sawabe, T.; Inazawa, J.; Tsuruo, T.
Oncogene 10, 1731-1738, 1995
A;Title: Cloning, expression and chromosomal localization of a novel gene for protein ty
A;Reference number: S60613; MUID:95273089; PMID:7753550
A;Accession: S60613
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1216 <SEI>
A;Cross-references: UNIPROT:Q16827; UNIPARC:UPI0000073BF3; EMBL:Z48541; NID:G963058; PID:
C;Superfamily: protein-tyrosine-phosphatase, receptor type O; fibronectin type III repea
C;Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatas
F;529-626/Domain: fibronectin type III repeat homology <3PR>
F;962-1184/Domain: protein-tyrosine-phosphatase homology <PRP1>
F;1136/Active site: Cys (phosphocysteine intermediate) #status predicted
F;1142/Binding site: substrate phosphate (Arg) #status predicted
```

```
Query Match 8.6% Score 906.5; DB 2; Length 1216;
Best Local Similarity 25.0%; Pred. No. 3.2e-42;
Matches 377; Conservative 190; Mismatches 459; Indels 484; Gaps 72;
```

```
Qy 655 GRNDYLSVSWL-VAPGDVDNYEVLTHDGVKQVSLVIKSVRECSFSSLPGRLYTVIT 713
Db 10 GARRLLPLLMLFVLKFNATAFHTVTVQDDNNIVVSL-----EAS-DVISPASVYVVKIT 61
Qy 714 TRSGKYENHSFSOER---TVPDKVQGVSVNSARSDDLRYSVSWHATGDFDHYETIKKN 770
Db 62 ---GESKNYFEFEENSTLPPV-----IFKASY-HGL-----YIITLVVN 101
Qy 771 NFITQTSIKPSKSENECFVQLVQGLYSVTVTKSGQYEAENEGNGRTIPBPVKDLTLNR 830
Db 102 GNVVTK-----PSR--SITVLT-----PLPVTISVSIYD 129
Qy 831 STE-----DLHVT-----MSGANGDV---DQVEIQLLFNDMKVFP----- 862
Db 130 KPSPEGTGLVEIHYPEKYNVFRVNIYSYMGKDFRTMLYKDFPKGKTVFNHW--LPQMCY 187
Qy 863 ---PHELV-----NTATEYRFTSLTPGROYKILVITISGDVQQSASFBSFTVPSAVKN 912
Db 188 SNTFQLVCBATNKSTVVEYSGVSHPE-KQHR-----TAPYPPQN 227
Qy 913 IHISPGATDSLTVNTPGGDVDSYTVSAFRHSQKVDSQTIPIKRVFEHTFHRLEAGEQY 972
Db 228 ISVRIYNLKN---NWEEQSG---NPEESFMSQ---DTICK---EKLHFTE--ETP 272
Qy 973 QIMIASVSGSLKQINQVNGRTVPASVQGVJADNAYSLSLVSWQAAQVAERYD----- 1027
Db 273 EIPSGNISSGWPD-----FNSSDYETTSQPYWWDASAPASEDEFFSV 316
Qy 1028 ILLLTENGILLRNTS-----EPATT----- 1047
Db 317 LPMEYENNSTLSETEKSTSGSFPPVQMLTWLPPKPPTAFDGFHIIHRENFTEYLM 376
Qy 1048 ---KQHKF--EDITPGKKYKIQILTV--SGGLFSKEAQTEGR---TVPAAVTDLRITEN 1096
Db 377 VDEEAHEFVAELKEPG-KYKLSVTTFTSSGSCETRSQSAKLSFYISPSGEWEELE-TE- 434
Qy 1097 STRHLGFR-----WTASEGELSWNFIYLNPDG-NLQERAQVDPLVQSFSPQ--- 1142
Db 435 KPQHVSHVLSSTTALMSWTSSQEN---YNSTIVSVVSLTCCQKESQRLKQYCTQVNS 491
Qy 1143 -----NLQGRMYKRWIYVTHSGEL---SNESIFIGRTVPASVSHLGRNSRNTTDSLWFN 1193
Db 492 SKPIIENLVPGAQYQVVIYLRKGLPIGPSPDPVTF-AIVPTGKIDMLVPLGPT-AVLVS 549
Qy 1194 WS-PASGDFDFY--ELILYNPNGTKKENW-----KDKDLT-EMRFQGLVPRKY---V 1239
Db 550 WTERPYLGVRKYVEMFYENP-ATWSEWTTYEIAATVSLTASVRIANLLPANYNFRV 608
Qy 1240 LWWVTHSGDLS--NKVTASRRTAPSPPLMSFADIANSLAITWKGPPDWTYNDPE-LQ 1296
Db 609 TWVTWGDPELSCDSDSTISFITAPVAPEITS-VEYFNLSLYISWTYGDGDTTDLSHSRMLH 667
```





GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2006, 18:19:33 ; Search time 193.96 Seconds  
(without alignments)  
4932.432 Million cell updates/sec

Title: US-10-633-742-6

Perfect score: 7095

Sequence: 1 MESKVLAVLWLCVETRAA.....GSTAQILQDPSTGLSSPPV 1356

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7092	99.9	1356	1 VGFPR2_HUMAN	P35968 homo sapien
2	7091	99.9	1451	2 Q59EB0_HUMAN	Q59EB0 homo sapien
3	6183.5	87.2	1343	2 Q5PQ00_RAT	Q5PQ00 rattus norv
4	6143.5	86.6	1343	1 VGFPR2_RAT	Q08775 rattus norv
5	6135.5	86.5	1345	2 Q8VCD0_MOUSE	Q8VCD0 mus musculu
6	6100.5	86.0	1367	1 VGFPR2_MOUSE	P35918 mus musculu
7	5091	71.8	1348	2 Q677M1_CHICK	Q677M1 gallus gall
8	5031.5	70.9	1348	1 VGFPR2_COTJA	P52583 coturnix co
9	3264.5	46.0	1357	2 Q5GIT4_BRARE	Q5GIT4 brachydanio
10	3058.5	43.1	1328	2 Q4RKW3_TETNG	Q4RKW3 tetraodon n
11	2981.5	42.0	1379	2 P79701_COTCO	P79701 coturnix co
12	2858.5	40.3	1357	2 Q5GIT2_BRARE	Q5GIT2 brachydanio
13	2855.5	40.2	1357	2 Q5MD89_BRARE	Q5MD89 brachydanio
14	2804	39.5	1363	1 VGFPR3_MOUSE	P35917 mus musculu
15	2804	39.5	1363	2 Q5SU94_MOUSE	Q5SU94 mus musculu
16	2790	39.3	1363	2 Q91ZT1_RAT	Q91ZT1 rattus norv
17	2751.5	38.8	1298	1 VGFPR3_HUMAN	P35916 homo sapien
18	2748.5	38.7	1338	1 VGFPR1_HUMAN	P17948 h vasculan
19	2746.5	38.7	1338	2 Q5TAK1_HUMAN	Q5TAK1 homo sapien
20	2742	38.6	1363	2 Q86W07_HUMAN	Q86W07 homo sapien
21	2734.5	38.5	1298	2 Q86W08_HUMAN	Q86W08 homo sapien
22	2677	37.7	1327	2 Q8QHL3_CHICK	Q8QHL3 gallus gall
23	2648	37.3	1336	1 VGFPR1_RAT	P53767 rattus norv
24	2631	37.1	1387	2 Q4SQ03_TETNG	Q4SQ03 tetraodon n
25	2623.5	37.0	1333	1 VGFPR1_MOUSE	P35969 mus musculu
26	2573.5	36.3	1302	1 VGFPR2_BRARE	Q8AXB3 brachydanio
27	2387	33.6	1272	2 Q5GIT3_BRARE	Q5GIT3 brachydanio
28	2376	33.5	1272	2 Q4JDD5_BRARE	Q4JDD5 brachydanio
29	2275.5	32.1	829	2 Q5T234_BRARE	Q5T234 brachydanio
30	2264.5	31.9	1092	2 Q91ZT0_RAT	Q91ZT0 rattus norv
31	2196.5	31.0	681	2 Q5DUF3_CHICK	Q5DUF3 gallus gall

32	1690.5	23.8	341	2 Q8CD05_MOUSE	Q8CD05 mus musculu
33	1634.5	23.0	464	2 Q5RIP2_BRARE	Q5RIP2 brachydanio
34	1621	22.8	458	2 Q5MD88_BRARE	Q5MD88 brachydanio
35	1583.5	22.3	551	2 Q59G09_HUMAN	Q59G09 homo sapien
36	1581.5	22.3	487	2 Q75WK4_ORYLA	Q75WK4 oryzias lat
37	1558.5	22.0	468	2 Q4RP85_TETNG	Q4RP85 tetraodon n
38	1408.5	19.9	1331	2 Q4S572_TETNG	Q4S572 tetraodon n
39	1343	18.9	346	2 Q42570_XENLA	Q42570 xenopus lae
40	1218.5	17.2	1089	1 PGFRA_HUMAN	P16234 homo sapien
41	1213	17.1	817	2 Q4H2M8_CIOIN	Q4H2M8 ciona intes
42	1202.5	16.9	1079	2 Q6E7G6_CANFA	Q6E7G6 canis famil
43	1201.5	16.9	1062	2 Q8AXC7_FUGRU	Q8AXC7 fugu rubrip
44	1192	16.8	1089	1 PGFRA_MOUSE	P26618 mus musculu
45	1192	16.8	1089	2 Q7TSJ3_MOUSE	Q7TSJ3 mus musculu

#### ALIGNMENTS

RESULT 1  
ID VGFPR2\_HUMAN STANDARD; PRT; 1356 AA.  
AC P35968; O60723; Q14178;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Vascular endothelial growth factor receptor 2 precursor (EC 2.7.1.112)  
DE (VEGFR-2) (Kinase insert domain receptor) (Protein-tyrosine kinase  
DE receptor Flk-1).  
GN Name-KDR; Synonyms=FLK1;  
OS Homo sapiens (Human)  
OC Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Yin L.Y., Wu Y., Patterson C.;  
RT "Full length human KDR/flk-1 sequence.";  
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RA Yu Y., Whitney R.G., Sato J.D.;  
RT "Coding region for human VEGF receptor KDR (VEGFR-2).";  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP NUCLEOTIDE SEQUENCE OF 3-1356.  
RC TISSUE=Umbilical vein;  
RX MEDLINE=92019839; PubMed=1656371;  
RA Terman B.I., Carrion M.E., Kovacs E., Rasmussen B.A., Eddy R.L.,  
RA Shows T.B.;  
RT "Identification of a new endothelial cell growth factor receptor  
RT tyrosine kinase.";  
RL Oncogene 6:1677-1683(1991).  
RN [4]  
RP NUCLEOTIDE SEQUENCE OF 1-22.  
RX MEDLINE=96032749; PubMed=7559454; DOI=10.1074/jbc.270.39.23111;  
RA Patterson C., Perrella M.A., Hsieh C.-M., Yoshizumi M., Lee M.-E.,  
RA Harber E.;  
RT "Cloning and functional analysis of the promoter for KDR/flk-1, a  
RT receptor for vascular endothelial growth factor.";  
RL J. Biol. Chem. 270:23111-23118(1995).  
RN [5]  
RP FUNCTION  
RX MEDLINE=93038639; PubMed=1417831;  
RA Terman B.I., Dougher-Vermazen M., Carrion M.E., Dimitrov D.,  
RA Armellino D.C., Gospodarowicz D., Boehlen P.;  
RT "Identification of the KDR tyrosine kinase as a receptor for vascular  
RT endothelial cell growth factor.";  
RL Biochem. Biophys. Res. Commun. 187:1579-1586(1992).  
CC -!- FUNCTION: Receptor for VEGF or VEGFR. Has a tyrosine-protein  
CC Kinase activity. The VEGF-kinase ligand/receptor signaling system

plays a key role in vascular development and regulation of vascular permeability.  
 -|- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + a protein tyrosine phosphate.  
 -|- SUBCELLULAR LOCATION: Type I membrane protein.  
 -|- SIMILARITY: Belongs to the tyr protein kinase family. CSP-1/PDGF receptor subfamily.  
 -|- SIMILARITY: Contains 7 Ig-like C2-type (immunoglobulin-like) domains.  
 -----  
 This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.  
 -----  
 DR EMBL; AF035121; AAB88005.1; -; mRNA.  
 DR EMBL; AF063658; AAC16450.1; -; mRNA.  
 DR EMBL; X61656; CAA43837.1; -; mRNA.  
 DR EMBL; L04947; AAA59459.1; -; mRNA.  
 DR EMBL; X89776; CAA61916.1; -; Genomic\_DNA.  
 DR PUB; JCI1402; JCI1402.  
 DR PUB; LYR2; X-ray; A=806-1171.  
 DR Ensembl; ENSG00000128052; Homo sapiens.  
 DR HGNC; HGNC:6307; KDR.  
 DR MIM; 191306; -.  
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
 DR GO; GO:0005021; F:vascular endothelial growth factor receptor. .; TAS.  
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; TAS.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR001824; RecepttyrkinsIII.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR InterPro; IPR008266; Tyr\_kinase\_AS.  
 DR InterPro; IPR009134; VEGFR.  
 DR InterPro; IPR009136; VEGFR2.  
 DR PRINTS; PR01832; VEGFRECEPTOR.  
 DR PRINTS; PR01834; VEGFRECEPTR2.  
 DR ProDom; PD000001; Prot\_kinase; 2.  
 DR SMART; SM00408; IGC2; 2.  
 DR SMART; SM00219; TyrcK; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 5.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 KW 3D-structure; Angiogenesis; ATP-binding; Developmental protein;  
 KW Differentiation; Glycoprotein; Immunoglobulin domain; Kinase;  
 KW Nucleotide-binding; Phosphorylation; Polymorphism; Receptor; Repeat;  
 KW Signal; Transference; Transmembrane; Tyrosine-protein kinase.  
 FT SIGNAL 1 19 Potential.  
 FT CHAIN 20 1356 Vascular endothelial growth factor receptor 2.  
 FT TOPO\_DOM 20 764 Extracellular (Potential).  
 FT TRANSMEM 765 789 Potential.  
 FT TOPO\_DOM 790 1356 Cytoplasmic (Potential).  
 FT DOMAIN 46 110 Ig-like C2-type 1.  
 FT DOMAIN 141 207 Ig-like C2-type 2.  
 FT DOMAIN 224 320 Ig-like C2-type 3.  
 FT DOMAIN 328 414 Ig-like C2-type 4.  
 FT DOMAIN 421 548 Ig-like C2-type 5.  
 FT DOMAIN 551 660 Ig-like C2-type 6.  
 FT DOMAIN 667 753 Ig-like C2-type 7.  
 FT DOMAIN 834 1162 Protein kinase.  
 FT NP\_BIND 840 848 ATP (By similarity).  
 FT ACT\_SITE 1028 1028 By similarity.  
 FT BINDING 868 868 ATP (By similarity).  
 FT MOD\_RES 1059 1059 Phosphotyrosine (by autocatalysis) (By similarity).  
 FT CARBOHYD 46 46 N-linked (GlcNAc. .) (Potential).  
 FT CARBOHYD 66 66 N-linked (GlcNAc. .) (Potential).  
 FT CARBOHYD 96 96 N-linked (GlcNAc. .) (Potential).

FT	CARBOHYD	143	143	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	158	158	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	245	245	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	318	318	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	374	374	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	395	395	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	511	511	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	523	523	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	580	580	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	613	613	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	619	619	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	631	631	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	675	675	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	704	704	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	721	721	N-linked (GlcNAc. .) (Potential).
FT	VARIANT	297	297	V -> I (in dbSNP:2305948).
FT	VARIANT	472	472	Q -> H (in dbSNP:1870377).
FT	CONFLICT	2	2	/FTId=VAR_020353.
FT	CONFLICT	772	772	Q -> E (in Ref. 2).
FT	CONFLICT	787	787	A -> T (in Ref. 3).
FT	CONFLICT	835	835	R -> G (in Ref. 3).
FT	CONFLICT	848	848	K -> N (in Ref. 3).
FT	CONFLICT	1347	1347	V -> E (in Ref. 3).
FT	CONFLICT	1347	1347	S -> T (in Ref. 3).
FT	SEQUENCE	1356	1356	AA; 151527 MW; 59E7C44B05CFEBB3 CRC64;

Query Match 99.9%; Score 7092; DB 1; Length 1356;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY	1	MESKULLAVALMVCVETRAASVGLPSVLDLRLSLQDKDILTIKANTTLQITCRGQRLD	60
DB	1	MQSKULLAVALMVCVETRAASVGLPSVLDLRLSLQDKDILTIKANTTLQITCRGQRLD	60
QY	61	WLWPNQSGSEQREVEVTECSGLFCFKTLTPKVIGNDTGAYCFYRETDLASVYVYVD	120
DB	61	WLWPNQSGSEQREVEVTECSGLFCFKTLTPKVIGNDTGAYCFYRETDLASVYVYVD	120
QY	121	YRSPFIASVSDQGVVYITENKNTVTPICLSINLNVSLCARYPEKFPDGNRISWD	180
DB	121	YRSPFIASVSDQGVVYITENKNTVTPICLSINLNVSLCARYPEKFPDGNRISWD	180
QY	181	SKKGFIPSYMISYAGWVCEAKINDESYQSIWYVWVGYRIYDVLSPSHGIELSVGE	240
DB	181	SKKGFIPSYMISYAGWVCEAKINDESYQSIWYVWVGYRIYDVLSPSHGIELSVGE	240
QY	241	KLVLNCTARTELNVGIDFNWEYPSKQHKLVNRDLKTQSGSEMKKFLSTLTIDGVTRS	300
DB	241	KLVLNCTARTELNVGIDFNWEYPSKQHKLVNRDLKTQSGSEMKKFLSTLTIDGVTRS	300
QY	301	DQGLYTCASSGLMTKKNSTFVRVHEKPFVAFGSGMESLVEATVGERVRIPAKYLGYPPP	360
DB	301	DQGLYTCASSGLMTKKNSTFVRVHEKPFVAFGSGMESLVEATVGERVRIPAKYLGYPPP	360
QY	361	EIKWYKNGIPLSNHTIKAGHVLTITIMEVSEDTGNVTILTNPISEKQSHVSVLVYVP	420
DB	361	EIKWYKNGIPLSNHTIKAGHVLTITIMEVSEDTGNVTILTNPISEKQSHVSVLVYVP	420
QY	421	PIQSKSLISPVDSYQYGTQTCTVVAIPPHPHIHWYQLEECANPPSAVSTNPFY	480
DB	421	PIQSKSLISPVDSYQYGTQTCTVVAIPPHPHIHWYQLEECANPPSAVSTNPFY	480
QY	481	PCEEVRSVEDFOGKNKIEVNKNQFALIEGKNKTSTLTVAQANVSALYKCEAVNKVGRGE	540
DB	481	PCEEVRSVEDFOGKNKIEVNKNQFALIEGKNKTSTLTVAQANVSALYKCEAVNKVGRGE	540
QY	541	RVISFHVTRGPITILQPDMPQTEQESVSLWCTADRTSTFENLTWYKLGPOPLPHVGELEPT	600
DB	541	RVISFHVTRGPITILQPDMPQTEQESVSLWCTADRTSTFENLTWYKLGPOPLPHVGELEPT	600
QY	601	PVCKNLDLTWLKLNATWPNSTNDIILIMELKNASLQDQGVYCLAODRKTGKHCVVRQLT	660

Db 601 PVCKNLDTLWKLNATMFSNSTNDILIMELKNASIQDQGDYVCLAQDRKTKKRHCVRQLT 660  
Qy 661 VLVRVAPTITGNLENQTTSGESIEVSCASGNPPPOIMWFKONETLVEDSGIVLKDGNR 720  
Db 661 VLVRVAPTITGNLENQTTSGESIEVSCASGNPPPOIMWFKONETLVEDSGIVLKDGNR 720  
Qy 721 NLTIIRVRKDEGLYTCQACSVLGCACVAFIIEGAQEKTNLEIILVGTAVIAMPFWL 780  
Db 721 NLTIIRVRKDEGLYTCQACSVLGCACVAFIIEGAQEKTNLEIILVGTAVIAMPFWL 780  
Qy 781 LLVILIRTVKRVANGGELKTYLSIVMDPDELPLDEHCHERLPLDASKWEPFRDRKLKGLKPL 840  
Db 781 LLVILIRTVKRVANGGELKTYLSIVMDPDELPLDEHCHERLPLDASKWEPFRDRKLKGLKPL 840  
Qy 841 GRGAFQGVIEADAFIDKATCRTAVAKMLKEGATHSEHRLMSELKILIHGHILNVN 900  
Db 841 GRGAFQGVIEADAFIDKATCRTAVAKMLKEGATHSEHRLMSELKILIHGHILNVN 900  
Qy 901 LLGACTKPGGLMWIVVEFCFKNLSTYLRSKRNEFVYKTKGARFRQGDYVGAIPVDLK 960  
Db 901 LLGACTKPGGLMWIVVEFCFKNLSTYLRSKRNEFVYKTKGARFRQGDYVGAIPVDLK 960  
Qy 961 RRLDSTSSOSSASSGFEVSEKSLSDVEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
Db 961 RRLDSTSSOSSASSGFEVSEKSLSDVEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
Qy 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR 1080  
Db 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR 1080  
Qy 1081 VYTIQSDVWFGVLLWEIPLSGASPYGVKIDEEFCRLKEGTRMAPDYTTPEMYQTM 1140  
Db 1081 VYTIQSDVWFGVLLWEIPLSGASPYGVKIDEEFCRLKEGTRMAPDYTTPEMYQTM 1140  
Qy 1141 DCWHGBSQRPPTSSELVEHLNLLQANAQDQDYIIVLPISSETLSMEEDGSLPTSPVS 1200  
Db 1141 DCWHGBSQRPPTSSELVEHLNLLQANAQDQDYIIVLPISSETLSMEEDGSLPTSPVS 1200  
Qy 1201 CMEEVEVCDPKFHYDNTAGISQVLSQNSKRKSRPVSVKTFEDIPLLEEPEVKVIPPDDNQDTS 1260  
Db 1201 CMEEVEVCDPKFHYDNTAGISQVLSQNSKRKSRPVSVKTFEDIPLLEEPEVKVIPPDDNQDTS 1260  
Qy 1261 GMVLASELKTLEDRTKLSFGWPSKRSVASEGNSQTSQYSGYHSDDTDTTVYS 1320  
Db 1261 GMVLASELKTLEDRTKLSFGWPSKRSVASEGNSQTSQYSGYHSDDTDTTVYS 1320  
Qy 1321 SERAELKLIEIGVGTGTAQILQPDGTTLSPPV 1356  
Db 1321 SERAELKLIEIGVGTGTAQILQPDGTTLSPPV 1356

## RESULT 2

Q59EB0\_HUMAN  
ID Q59EB0\_HUMAN PRELIMINARY; PRT; 1451 AA.  
AC Q59EB0;  
DT 10-MAY-2005 (TrEMBLrel. 30, Created)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)  
DE Kinase insert domain receptor (A type III receptor tyrosine kinase)  
DE variant (fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Aorta endothelial cell;  
RA Toki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,  
RA Ohara O., Nagase T., Kikuno F.R.;  
RT "None Title."  
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB209901; BAD93138.1; -; mRNA.

KW Kinase; Receptor.  
FT NON\_TER 1  
SQ SEQUENCE 1451 AA; 161601 MW; 915FB64B51EFA0CB CRC64;  
Query Match 99.9%; Score 7091; DB 2; Length 1451;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1354; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MESKVLALVALMVCVETRAASVGLPSVSLDPLRLSTQKQILTIKANTTLOITCRGORDLD 60  
Db 96 MQSKVLALVALMVCVETRAASVGLPSVSLDPLRLSTQKQILTIKANTTLOITCRGORDLD 155  
Qy 61 WLWPNNSQSEQRVEVTECSGDLFCCKTLTIKPVIGNDTGAYKCFYRETDLASVIVYVQD 120  
Db 156 WLWPNNSQSEQRVEVTECSGDLFCCKTLTIKPVIGNDTGAYKCFYRETDLASVIVYVQD 215  
Qy 121 YRSPFIASVSDQGVVYITENKNTVIVPCLSISNLNVSLCARYPEKRPVDPGNRISWD 180  
Db 216 YRSPFIASVSDQGVVYITENKNTVIVPCLSISNLNVSLCARYPEKRPVDPGNRISWD 275  
Qy 181 SKKGFTIPSYMISYAGWVCEAKINDESQSYIMYIVVVGRIYDVVLSPSHGIELSVGE 240  
Db 276 SKKGFTIPSYMISYAGWVCEAKINDESQSYIMYIVVVGRIYDVVLSPSHGIELSVGE 335  
Qy 241 KLVLNCTARTELNVGIDFNWEYPSKHQKLVNRDLKTQSGSEMKEFLSTLTIDGVTRS 300  
Db 336 KLVLNCTARTELNVGIDFNWEYPSKHQKLVNRDLKTQSGSEMKEFLSTLTIDGVTRS 395  
Qy 301 DQGLYTCASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGBRVIRPAKYLGYPPP 360  
Db 396 DQGLYTCASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGBRVIRPAKYLGYPPP 455  
Qy 361 EIKWYKNGIPILESNHTIKAGHVTIMEVSRDGTNTVILTNPISEKQSHVSVLVYVVP 420  
Db 456 EIKWYKNGIPILESNHTIKAGHVTIMEVSRDGTNTVILTNPISEKQSHVSVLVYVVP 515  
Qy 421 PQIGESLISPDVSYQYGTQTLCTVYAIIPPHIHWYQLEECANESQAVSVTNPY 480  
Db 516 PQIGESLISPDVSYQYGTQTLCTVYAIIPPHIHWYQLEECANESQAVSVTNPY 575  
Qy 481 PCBEWSVEDFOGCKIEVKNQOFALIEGKNTVSTLVIQAAVNSALYKCEAVNKVGRGE 540  
Db 576 PCBEWSVEDFOGCKIEVKNQOFALIEGKNTVSTLVIQAAVNSALYKCEAVNKVGRGE 635  
Qy 541 RVISFHVTRGPEITLQPDQPTQESVSLWCTADRSTFENLTWYKLGQPLPIHVGELPT 600  
Db 636 RVISFHVTRGPEITLQPDQPTQESVSLWCTADRSTFENLTWYKLGQPLPIHVGELPT 695  
Qy 601 PVCKNLDTLWKLNATMFSNSTNDILIMELKNASLQDQGDYVCLAQDRKTKKRHCVRQLT 660  
Db 696 PVCKNLDTLWKLNATMFSNSTNDILIMELKNASLQDQGDYVCLAQDRKTKKRHCVRQLT 755  
Qy 661 VLVRVAPTITGNLENQTTSGESIEVSCASGNPPPOIMWFKONETLVEDSGIVLKDGNR 720  
Db 756 VLVRVAPTITGNLENQTTSGESIEVSCASGNPPPOIMWFKONETLVEDSGIVLKDGNR 815  
Qy 721 NLTIIRVRKDEGLYTCQACSVLGCACVAFIIEGAQEKTNLEIILVGTAVIAMPFWL 780  
Db 816 NLTIIRVRKDEGLYTCQACSVLGCACVAFIIEGAQEKTNLEIILVGTAVIAMPFWL 875  
Qy 781 LLVILIRTVKRVANGGELKTYLSIVMDPDELPLDEHCHERLPLDASKWEPFRDRKLKGLKPL 840  
Db 876 LLVILIRTVKRVANGGELKTYLSIVMDPDELPLDEHCHERLPLDASKWEPFRDRKLKGLKPL 935  
Qy 841 GRGAFQGVIEADAFIDKATCRTAVAKMLKEGATHSEHRLMSELKILIHGHILNVN 900  
Db 936 GRGAFQGVIEADAFIDKATCRTAVAKMLKEGATHSEHRLMSELKILIHGHILNVN 995  
Qy 901 LLGACTKPGGLMWIVVEFCFKNLSTYLRSKRNEFVYKTKGARFRQGDYVGAIPVDLK 960  
Db 996 LLGACTKPGGLMWIVVEFCFKNLSTYLRSKRNEFVYKTKGARFRQGDYVGAIPVDLK 1055  
Qy 961 RRLDSTSSOSSASSGFEVSEKSLSDVEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020

```

Db      1056 RRLDSTSSQSSASSGVEEKSLSDEEEAPEDLYKDFLTLEHLICYSQVAKGMEFLA 1115
QY      1021 SRKCIHRDLAARNILLSEKNVKKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR 1080
Db      1116 SRKCIHRDLAARNILLSEKNVKKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR 1175
QY      1081 VYTIQSDVWSFGVLLWEIFSLGASPYGVKIDBEFCRRLLKEGTRMRAPDYTTPEMYQTML 1140
Db      1176 VYTIQSDVWSFGVLLWEIFSLGASPYGVKIDBEFCRRLLKEGTRMRAPDYTTPEMYQTML 1235
QY      1141 DCWHGSPSQPTSELSVEHLGNLLQANAQDGKDIYVLPISSETLSNEEDSGLSPTSPVS 1200
Db      1236 DCWHGSPSQPTSELSVEHLGNLLQANAQDGKDIYVLPISSETLSNEEDSGLSPTSPVS 1295
QY      1201 CMEEEEVCDPKFYDNTAGISQYQLNSKRKSRPVSVKTFEDIPLEBPVVKVIPPDDNOTDS 1260
Db      1296 CMEEEEVCDPKFYDNTAGISQYQLNSKRKSRPVSVKTFEDIPLEBPVVKVIPPDDNOTDS 1355
QY      1261 GMYLASEELKLEDRKTLSPFSGMVPKSRRESVASEGNSQTSYGQSYGSHSDDTDTTVYS 1320
Db      1356 GMYLASEELKLEDRKTLSPFSGMVPKSRRESVASEGNSQTSYGQSYGSHSDDTDTTVYS 1415
QY      1321 SEBAELLKLEIGVQTGSTAQIILQPSDGTTLSPPV 1356
Db      1416 SEBAELLKLEIGVQTGSTAQIILQPSDGTTLSPPV 1451

RESULT 3
QSFQQUO_RAT PRELIMINARY; PRT; 1343 AA.
AC QSFQQUO;
DT 01-FEB-2005 (T-EMBLrel. 29, Created)
DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (T-EMBLrel. 29, Last annotation update)
DE Kdr protein.
GN Name=kdr;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vialalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lung;
RA Director MGC Project;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC087029; AAB87029.1; -; mRNA.
GO; GO:0016020; C:membrane; IEA.

```

```

DR GO:0005524; F:ATP binding; IEA.
DR GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO:0005021; F:vascular endothelial growth factor receptor. .; IEA.
DR GO:000468; P:protein amino acid phosphorylation; IEA.
DR GO:0007169; P:transmembrane receptor protein tyrosine kin. .; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001824; RecepttyrkinaIII.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR InterPro; IPR009134; VEGFR.
DR InterPro; IPR009136; VEGFR2.
DR Pfam; PF00047; ig_1.
DR PRINTS; PR01832; VEGFRECEPTOR.
DR PRINTS; PR01834; VEGFRECEPTR2.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00409; IG; 7.
DR SMART; SM00408; IGC2; 5.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00835; IG-LIKE; 5.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TVR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
SQ SEQUENCE 1343 AA; 150274 MW; F4B906E8012A5C59 CRC64;

Query Match 87.2%; Score 6183.5; DB 2; Length 1343;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 1170; Conservative 75; Mismatches 98; Indels 13; Gaps 2;

QY 1 MESKLLAVALWLCVETRAASVGLPSVSLDLPRLSLQKDLITKANTTLQITCRGORDLD 60
DB 1 MESKLLAVALWLCVETRAASVGLPSVSLDLPRLSLQKDLITKANTTLQITCRGORDLD 60

QY 61 WLWPNQSGSEORVEVTECSGLFCCKTLTPKVIGNDTGAYKCFYRETDLASIVVYVVD 120
DB 61 WLWPNTPRDESERVLTECGDSIFCKTLTPRVGVNDTGAYKCFYRDTDVSIYVYVVD 120

QY 121 YRSPFIASVSDQGVVYITENKNTKTVIPCLGSIENLVNLSLCARYPEKRFVDPGNRISWD 180
DB 121 HRSPFIASVSDHEGIVYITENKNTKTVIPCRGSIENLVNLSLCARYPEKRFVDPGNRISWD 180

QY 181 SKGFTIPSYMISYAGMVFCEAKINDESVQSYTMVIVVGVYRIYDVLSPSHGIELSVCE 240
DB 181 SEKGTIPSYMISYAGMVFCEAKINDESVQSYTMVIVLVVGVYRIYDVLSPPHEIELSAGE 240

QY 241 KLVLNCTARTELNVGIDFNWEYPSKSHQKHLVNRDLKTSQSGSEMKKFLSTLTIDGVTRS 300
DB 241 KLVLNCTARTELNVGLDFSMQFPSSKHQKHLVNRDVKSLPGTVAKMFLSTLTIDSVTKS 300

QY 301 DGLYLTCAASSGLMTKQNSTFVRVHEKFPVAPGSGMESLVEATVGGERVPAIKYLYGPPP 360
DB 301 DQGEYCTAYSGLMTKQNKTFVRVHTKFPFAPGSGMKSIVEATVGSQVRIPVKYLSYPAP 360

QY 361 EIKWYKNGIPLESNHTIKAGHVLTITMEYSERDTGNVTTLTNPISEKESHVSVLVVYYP 420
DB 361 DIKYRNGRPIESNHTIMVIGDELITMEYSERDAGNTVTLTNPISEKESHVSVLVVYYP 420

QY 421 PQIGKSLISPVDSYQYGTQTLCTCTVVAIPPHPHIHWYQLEEEECANEPSPQAVSVTNPY 480
DB 421 PQIGKALISPMDSYQYGTMTQTLCTCTVYANPLPHIOWYQLEEACSVPSQ----TNPY 476

QY 481 PCEEWSVEDFQGGNKIEVNKNQFALIEGKNKTVSTLVIQAANVSALYKCEAVNKVGRGE 540
DB 477 TCKEWRHVKDFQGGNKIEVTKNQYALIEGKNKTVSTLVIQAANVSALYKCEAINKAGRGE 536

QY 541 RVISFHVTRGPETLQPDMDQTEQESVSLWCTADSTRFENLTWYKLGQPLPIHVGELEPT 600
DB 537 RVISFHVIRGPETVQPAQTQTEQESVSLCTADRTFENLTWYKLGQATSVHMGESLT 596

```





FT	CARBOHYD	245	245	N-linked (GlcNAc. . .)	(Potential).
FT	CARBOHYD	318	318	N-linked (GlcNAc. . .)	(Potential).
FT	CARBOHYD	374	374	N-linked (GlcNAc. . .)	(Potential).
FT	CARBOHYD	395	395	N-linked (GlcNAc. . .)	(Potential).
FT	CARBOHYD	507	507	N-linked (GlcNAc. . .)	(Potential).
FT	CARBOHYD	576	576	N-linked (GlcNAc. . .)	(Potential).
FT	CARBOHYD	609	609	N-linked (GlcNAc. . .)	(Potential).
FT	CARBOHYD	615	615	N-linked (GlcNAc. . .)	(Potential).
FT	CARBOHYD	627	627	N-linked (GlcNAc. . .)	(Potential).
FT	CARBOHYD	671	671	N-linked (GlcNAc. . .)	(Potential).
FT	CARBOHYD	700	700	N-linked (GlcNAc. . .)	(Potential).
FT	CARBOHYD	717	717	N-linked (GlcNAc. . .)	(Potential).
SQ	SEQUENCE	1343	AA; 150394	NW; AD7E509EB62D3FF4	CRC64;
Query Match					
Best Local Similarity 86.6%; Score 6143.5; DB 1; Length 1343;					
Matches 1161; Conservative 80; Mismatches 102; Indels 13; Gaps 2;					
Qy	1	MESKVLAVALLWC	VETRAASVGLPSVSLDPLRLSIQKQILITIKANTTILQITCRGQDRLD	60	
Db	1	MESRALLAVALLWFC	VETRAASVGLPGDSLHPPKLSQKQILITILANTTILQITCRGQDRLD	60	
Qy	61	WLFPNQSGSEQRVEVTECS	DGLFCKTLTIPKVIQNDTGAYKCFYRETDLASVIYVYQD	120	
Db	61	WLFPNTPRDSEERVLVTE	CGDSIFCKTLTPVRVVGNDTGAYKCFYRTDVSIVYVYQD	120	
Qy	121	YRSPFTASVSDQGVVYIT	ENKNTVVICPLGSGISNLNVSLCARYPEKFRVDPGNRISWD	180	
Db	121	HRSPFTASVSDERGI	VITENKNTVVICPGSGISNLNVSLCARYPEKFRVDPGNRISWD	180	
Qy	181	SKSGFTIPSMISYAGV	MVCEAKINDESQSIIMYIVVVGRIYDVVLSPSHGIELSVGE	240	
Db	181	SEKGFPTIPSMISYAG	VVCEAKINDETYQSIIMYIVLVGRIYDVVLSPPHIEIELSAGE	240	
Qy	241	KLVLNCTARTELNVG	IDFWNEYPSSKHQHKUVRNOLKTQSGSEMKKFSTLTIDGVTRS	300	
Db	241	KLVLNCTARTELNVGL	DFSMQFPSSKHQHKKIVNRDVKLPSPGTVAKMFSTLTIDSVTKS	300	
Qy	301	DOGLYTCASSGLMTK	KNSTFVRVHEKPFVAFGSGMESLVEATVGRVRIPAKYLGYPYP	360	
Db	301	DOGEYCTAYSGLM	TAKNTKTVRVTKPFATFGSGMKSLEATVGSQVRIPVKILSYAP	360	
Qy	361	EIKWYKNGIPLSNHT	IKAGHVLITIMEVSESDTGNVTILTNIPISKEKQSHVVSVLVVYP	420	
Db	361	DIKWYRNGRIPIES	NYTMIVGDELTIMEVSESDAGNTVILTNIPISMEKQSHVVSVLVVNP	420	
Qy	421	PQIGKSLISPVDSY	QYGTQTTLCTVYALPPHHIHWLQEBEACANEPSSQAVSTNYP	480	
Db	421	PQIGKALISPMDSY	QYGTMTLCTVYANPLHHIQWYWLQEBEACSYRPSQ----TNYP	476	
Qy	481	PCBEWSEVDFQGN	KIEVKNQFALIEGKNKTVSLVIOANVVSALYKCEAVNKVGRGE	540	
Db	477	TKWEHRVDFQGN	KIEVTKQYALIEGKNKTVSLVIOAYVSALYKCEALNKAGRGE	536	
Qy	541	RVISFHVTRGPET	ITLQPDQPTQEQESVSLWCTADRTFFENLTWYKLGQPLPHVGE	600	
Db	537	RVISFHVIRGPET	ITVQATQPTERESMSLLCTADRTFFENLTWYKLGQATSVHGESLT	596	
Qy	601	PVCNKLDITLWKN	ATMFSNSTDILIMELKNASLOQDQVYVLAQDRKTKGRHVVRQUT	660	
Db	597	PVCNKLDALWKL	NGTVFSNSTDILIVAFONASLOQDQNVYCSAQDKTKGRHCLVQLV	656	
Qy	661	VLERVAPTITGN	LENQTTISGESIEVYSCASGNPPQIMWFKDNFTLVSDSGIVLKDGR	720	
Db	657	ILERMAMPITGN	LENQTTITIGETIEVVCPTSGNPTPLITWFKDNFTLVSDSGIVLKDGR	716	
Qy	721	NLTIRVRKDEGLY	TCQCSVLGCACVBAFFIIEGAQEKTNLEIILVGTAVIAMFFWL	780	
Db	717	NLTIRVRKDEGLY	TCQCNVLGCARAEFTLIEIGVQEKTNLEVLVGTAVIAMFFWL	776	
Qy	781	LLVILIRTVKR	ANGBELKTYGISVMDPDELPLDEHCERLPHYDASKWEPFRDLKLGKPL	840	
Db	777	LLVILIRTVKR	ANGBELKTYGISVMDPDELPLDERCERLPHYDASKWEPFRDLKLGKPL	836	

RESULT 5	
Q8VCD0_MOUSE	
ID	Q8VCD0_MOUSE PRELIMINARY; PRT; 1345 AA.
AC	Q8VCD0;
DT	01-MAR-2002 (TREMBLrel. 20, Created)
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE	Kinase insert domain protein receptor.
GN	Name=Kdr;
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC	Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=FVB/N; TISSUE=Kidney;
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA	Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Villalón D.K., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA	Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]

RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=FVB/N; TISSUE=Kidney;  
 RA Director MGC Project;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC020530; AH20530.1; -, mRNA.  
 DR HSP; P35968; 1VR2.

DR Ensemble; ENSMUSG0000062960; Mus musculus.

DR MGI; MGI:96683; Kdr.

DR GO; GO:0005615; C:extracellular space; TAS.

DR GO; GO:0045165; P:cell fate commitment; IMP.

DR GO; GO:0016477; P:cell migration; IGR.

DR GO; GO:0045446; P:endothelial cell differentiation; IDA.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003598; Ig\_c2.

DR InterPro; IPR000719; Prot kinase.

DR InterPro; IPR001824; Receptor tyrosine kinase III.

DR InterPro; IPR001245; Tyr\_kinase.

DR InterPro; IPR008266; Tyr\_kinase\_AS.

DR InterPro; IPR009134; VEGFR.

DR Pfam; PF00047; Ig\_1.

DR PRINTS; PR01832; VEGFR-RECEPTOR.

DR PRINTS; PR01834; VEGFR-RECEPTOR2.

DR ProDom; PD000001; Prot kinase; 2.

DR SMART; SM00408; IgC2; 1.

DR SMART; SM00219; Tyrc; 1.

DR PROSITE; PS00835; IG LIKE; 5.

DR PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN\_1.

DR PROSITE; PS00011; PROTEIN KINASE\_DOM; 1.

DR PROSITE; PS00109; PROTEIN KINASE\_TYR; 1.

DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.

KW Kinase; Receptor.

SQ SEQUENCE 1345 AA; 150460 MW; 11859F8A58A33A39 CRC64;

Query Match 86.5%; Score 6135.5; DB 2; Length 1345;

Best Local Similarity 85.9%; Pred. No. 0;

Matches 1166; Conservative 71; Mismatches 106; Indels 15; Gaps 3;

QY 1 MESKALLAVALMVCVETRAASVGLPSVSLDPLRLSIQKDLITLTKANTLLOITCRGORDLD 60  
 DB 1 MESKALLAVALMFCVETRAASVGLPGLPPLPKLSTOKDILTILANTLLOITCRGORDLD 60  
 QY 61 WLFPNNSQSGEQRVEVTEC--SDGLFCKTLTIPKVIQNDTGAYKCFYRETDLASVIYVYV 118  
 DB 61 WLFPNAQRDSEERVLVTECGGDSIFCKTLTIPRVVGNDTGAYKCSYRDVDIASVYVYV 120  
 QY 119 QYRSPPIASVSDQGVVYITENKNTWVTPCLSGISNLNLSLCARYPEKRFVPDGNRIS 178  
 DB 121 RYRSPPIASVSDQGVVYITENKNTWVTPCRGISNLNLSLCARYPEKRFVPDGNRIS 180  
 QY 179 WDSKKGITPSYMSIYAGMVFCEAKINDSEYQSTIMYIVVVGRIYDVVLSPSHGLTSLV 238  
 DB 181 WDSIEIGTLPYSYMSIYAGMVFCEAKINDEYQSTIMYIVVVGRIYDVVLSPSHGLTSLV 240  
 QY 239 GEKLVNCTARTELNVLGIDFNWEYPSKSHKHKLNVRLDKLTQSGSEMKFLSTLTIDGVT 298  
 DB 241 GEKLVNCTARTELNVLGIDFNWEYPSKSHKHKLNVRLDKLTQSGSEMKFLSTLTIESVT 300  
 QY 299 RSDQGLYTCAASGLMFKNTSTFVRVHEKDFVAFGSGMESLVEATVGERVIRIPAKYLYGP 358  
 DB 301 KSDQGYTCVASSGRMIKRNRTFVRVHTKPFIAFGSGMKSLVEATVGSQVRIPVKYLYSP 360  
 QY 359 PPEIKWYKNGIPLSNHTIKAGHVLTIMEYSEKDTGNYTILTNPISKEQSHVSLVYV 418  
 DB 361 APDIKWYRNGRPIESNYTMTVGDELTIMEYTERDAGNYTILTNPISMERQSHVSLVYV 420  
 QY 419 VPPQIGSKSLIPVDSYQYGTOTLTCTVVAIPPHHIIHWYQLEECANEPQSAVSVTN 478  
 DB 421 VPPQIGKALISPMDSYQYGTOTLTCTVVAIPPHHIIHWYQLEECACSTRPQ---TS 476

QY 479 PYCEEWRSEVEDQGGNKIEVKNOPALIEGKNKTSTLVIQAAVNSALYKCEAVNKVR 538  
 DB 477 PYACKWRHVEDQGGNKIEVKNQVALIEGKNKTSTLVIQAAVNSALYKCEAVNKVR 536  
 QY 539 GERVISFHVTRGPEITLQPDQMTEQESVSLWCTADRTSTPENLTWYKLGQPLPIHVGEL 598  
 DB 537 GERVISFHVTRGPEITVQPAQTEQESVSLCTADRTSTPENLTWYKLGQSAQSVHMGES 596  
 QY 599 PTPVCKNLDLTKLNATMFSNSTNDILIMELKNASLQDQGYVCLADRTKRGKHCVVVRQ 658  
 DB 597 LTPVCKNLDLTKLNATMFSNSTNDILIVAFNASLQDQGYVCSAQDKTKRKHCLVKQ 656  
 QY 659 LTVLERVAPITGNLENQNTSICESLEVSCTASGNPPQIMWFKNETLVEDSGIVLKDQ 718  
 DB 657 LILERWAPITGNLENQNTTIGETIEVTPASGNPPTHITWFKNETLVEDSGIVLRDG 716  
 QY 719 NRNLTIIRVRKEDGLYTCQACSVLGCACVKAEPFIIEGAQKTNLEIILVGTAVIAMFF 778  
 DB 717 NRNLTIIRVRKEDGLYTCQACNVLCARAEFIIEGAQKTNLEIILVGTAVIAMFF 776  
 QY 779 WLLLVILRTVKRANGELKTYLSIVMDPELPLDEHCERLKYDASKWFFPRDLKLGK 838  
 DB 777 WLLLVILRTVKRANGELKTYLSIVMDPELPLDEHCERLKYDASKWFFPRDLKLGK 836  
 QY 839 PLGRGAGQVIEADAFGIDKTATCTRTVAVKMLKEGATHSEHRAIMSELKILIHGHILNV 898  
 DB 837 PLGRGAGQVIEADAFGIDKTATCTRTVAVKMLKEGATHSEHRAIMSELKILIHGHILNV 896  
 QY 899 VNLGACTKPGGLMWIVVEFCFKNLSTYLRSKRNEFPVYKTKGARPROCKQDVYGAIPVD 958  
 DB 897 VNLGACTKPGGLMWIVVEFCFKNLSTYLGRNEFPVYKTKGARPROCKQDVYGAIPVD 956  
 QY 959 LKRRLDSITSSQSSASSGFFVEEKSLSDEVEEAPEDLYKDFLTLEHLICYSFQVAKGMEF 1018  
 DB 957 LKRRLDSITSSQSSASSGFFVEEKSLSDEVEEASEELYKDFLTLEHLICYSFQVAKGMEF 1016  
 QY 1019 LASRKCIRHDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKMAPETIF 1078  
 DB 1017 LASRKCIRHDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKMAPETIF 1076  
 QY 1079 DRVYTIQSDVMSVGLWELFSLGASYPGVKIDEEFCRLKEGTNRARADYTTPEMYQT 1138  
 DB 1077 DRVYTIQSDVMSVGLWELFSLGASYPGVKIDEEFCRLKEGTNRARADYTTPEMYQT 1136  
 QY 1139 MLDCEWHEBPSQRPTFSELVEHLGNLLQANAQDQKDYIVLPISETLSMEEDSGLSLTSP 1198  
 DB 1137 MLDCEWHEBPSQRPTFSELVEHLGNLLQANAQDQKDYIVLPISETLSMEEDSGLSLTSP 1196  
 QY 1199 VSCWHEEEVCDPKFHYDNTAGISQYLNQSKRKRSPVSVKTFEDIPLPEEPVKVIPPDDNQ 1258  
 DB 1197 VSCWHEEEVCDPKFHYDNTAGISQYLNQSKRKRSPVSVKTFEDIPLPEEPVKVIPPDDNQ 1256  
 QY 1259 DSGMVLASBELKTLDEORTKLSPSFGMWVPSKRESVASEGNSQTSYGQSYGSHSDDDTTTV 1318  
 DB 1257 DSGMVLASBELKTLDEORTKLSPSFGMWVPSKRESVASEGNSQTSYGQSYGSHSDDDTTTV 1316  
 QY 1319 YSSEAEELKLEIGVGTGTGTAQIQDPSGTTLSPPV 1356  
 DB 1317 YSSEAEELKLEIGVGTGTGTAQIQDPSGTTLSPPV 1354  
 RESULT 6  
 ID VGF2 MOUSE STANDARD; PRT; 1367 AA.  
 AC P35918;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 13-SEP-2005 (Rel. 48, Last annotation update)  
 DE Vascular endothelial growth factor receptor 2 precursor (EC 2.7.1.112)  
 DE (VEGFR-2) (Protein-tyrosine kinase receptor flk-1) (Fetal liver kinase  
 DE 1) (Kinase NYK).  
 GN Name=Kdr; Synonyms=Flk-1, Flk1;

OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridea; Muridae; Murinae; Mus.  
ON NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=BALB/c; TISSUE=Embryo;  
RX MEDLINE=93208880; PubMed=7681362; DOI=10.1016/0092-8674(93)90573-9;  
RA Mollauer B., Witzmann-Voos S., Schnurch H., Martinez R.,  
RA Mueller N.P.H., Rieau W., Ullrich A.;  
RT "High affinity VEGF binding and developmental expression suggest Flk-1  
RT as a major regulator of vasculogenesis and angiogenesis.";  
RL Cell 72:835-846(1993).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C3H/He; TISSUE=Fetal liver;  
RX MEDLINE=92020984; PubMed=1717995;  
RA Mathews W., Jordan C.T., Gavin M., Jenkins N.A., Copeland N.G.,  
RA Lemishcka I.R.;  
RT "A receptor tyrosine kinase cDNA isolated from a population of  
RT enriched primitive hematopoietic cells and exhibiting close genetic  
RT linkage to c-kit.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:9026-9030(1991).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC MEDLINE=93141255; PubMed=8423988;  
RX Oelrichs R.B., Reid H.H., Bernard O., Ziemiecki A., Wilks A.F.;  
RA "NYK/Flk-1: a putative receptor protein tyrosine kinase isolated from  
RT E10 embryonic neuroepithelium is expressed in endothelial cells of the  
RT developing embryo.";  
RL Oncogene 8:11-18(1993).  
RN [4]  
RP NUCLEOTIDE SEQUENCE OF 1-15.  
RX MEDLINE=96032749; PubMed=7559454; DOI=10.1074/jbc.270.39.23111;  
RA Patterson C., Ferrerella M.A., Hsieh C.-M., Yoshizumi M., Lee M.-E.,  
RA Harber E.;  
RT "Cloning and functional analysis of the promoter for KDR/flk-1, a  
RT receptor for vascular endothelial growth factor.";  
RL J. Biol. Chem. 270:23111-23118(1995).  
RN [5]  
RP FUNCTION.  
RX MEDLINE=93361481; PubMed=8356051;  
RA Quinn T.P., Peters K.G., de Vries C., Ferrara N., Williams L.T.;  
RT "Fetal liver kinase 1 is a receptor for vascular endothelial growth  
RT factor and is selectively expressed in vascular endothelium.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:7533-7537(1993).  
CC -1- FUNCTION: Receptor for VEGF or VEGFC. Has a tyrosine-protein  
CC kinase activity. The VEGF-kinase ligand/receptor signaling system  
CC plays a key role in vascular development and regulation of  
CC vascular permeability.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + a protein  
CC tyrosine phosphate.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- TISSUE SPECIFICITY: Expressed at high levels in adult heart, lung,  
CC kidney, brain and skeletal muscle, but is also expressed at lower  
CC levels in most other adult tissues.  
CC -1- SIMILARITY: Belongs to the Tyr protein kinase family. CSF-1/PDGF  
CC receptor subfamily.  
CC -1- SIMILARITY: Contains 7 Ig-like C2-type (immunoglobulin-like)  
CC domains.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
CC ENBL: X70842; CAA50192.1; -; mRNA.  
CC ENBL: X59397; CAA42040.1; -; mRNA.  
CC ENBL: SS3103; AAB25043.1; -; mRNA.  
CC ENBL: X89777; CAA61917.1; -; Genomic\_DNA.  
CC PIR: A41228; A41228.

DR HSSP; P35968; 1VR2.  
DR Ensembl; ENSMUSG00000062960; Mus musculus.  
DR MGI; MGI:96683; Kdr.  
DR GO; GO:0005615; C:extracellular space; TAS.  
DR GO; GO:0045165; P:cell fate commitment; IMP.  
DR GO; GO:0016477; P:cell migration; IGI.  
DR GO; GO:0045446; P:endothelial cell differentiation; IDA.  
DR GO; GO:0030097; P:hemopoiesis; IMP.  
DR GO; GO:0001570; P:vasculogenesis; IMP.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003598; Ig\_c2.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR001824; Recepttyr\_kinIII.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR InterPro; IPR008266; Tyr\_pkinase\_AS.  
DR InterPro; IPR009134; VEGFR.  
DR Pfam; PF00047; ig; 1.  
DR PRINTS; PR01834; VEGFRECEPTOR.  
DR PRINTS; PR01832; VEGFRECEPTOR.  
DR ProDom; PD000001; Prot\_kinase; 2.  
DR SMART; SM00408; IGC2; 1.  
DR SMART; SM00219; Tyrc; 1.  
DR PROSITE; PS00835; IG LIKE; 5.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
KW Angiogenesis; ATP-binding; Developmental protein; Differentiation;  
KW Glycoprotein; Immunoglobulin domain; Kinase; Nucleotide-binding;  
KW Phosphorylation; Receptor; Repeat; Signal; Transferase; Transmembrane;  
KW Tyrosine-protein kinase.  
FT SIGNAL 1 19 Potential.  
FT CHAIN 20 1367 Vascular endothelial growth factor  
FT receptor 2.  
FT Extracellular (Potential).  
FT Transmem 763 784 Potential.  
FT TOPO\_DOM 20 762 Cytoplasmic (Potential).  
FT TOPO\_DOM 785 1367 Ig-like C2-type 1.  
FT DOMAIN 46 111 Ig-like C2-type 2.  
FT DOMAIN 143 209 Ig-like C2-type 3.  
FT DOMAIN 226 325 Ig-like C2-type 4.  
FT DOMAIN 330 416 Ig-like C2-type 5.  
FT DOMAIN 423 542 Ig-like C2-type 6.  
FT DOMAIN 549 656 Ig-like C2-type 7.  
FT DOMAIN 665 751 Protein kinase.  
FT NP\_BIND 832 1160 ATP (By similarity).  
FT ACT\_SITE 1026 1026 By similarity.  
FT BINDING 866 866 ATP (By similarity).  
FT MOD\_RES 1057 1057 Phosphotyrosine (by autocatalysis) (By  
FT similarity).  
FT CARBOHYD 46 46 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 98 98 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 145 145 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 160 160 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 247 247 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 320 320 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 376 376 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 397 397 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 509 509 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 521 521 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 578 578 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 611 611 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 617 617 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 629 629 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 673 673 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 702 702 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 719 719 N-linked (GlcNAc...) (Potential).  
FT CONFLICT 25 25 P -> T (in Ref. 1).  
FT CONFLICT 679 679 LV -> VL (in Ref. 1).  
FT CONFLICT 783 784 S -> C (in Ref. 1).  
FT CONFLICT 917 917 QUTSLGSGPVPAPPTPGNHERGAA -> RSPVP (in  
FT CONFLICT 1341 1367

FT SQ SEQUENCE 1367 AA; 152517 MW; EFC99704FIDCA266 CRC64; Ref. 3).

Query Match 86.0%; Score 6100.5; DB 1; Length 1367;  
Best Local Similarity 85.7%; Pred. No. 0;  
Matches 1160; Conservative 72; Mismatches 106; Indels 15; Gaps 3;

Qy 1 MESKLLAVALWLCVETRAASVGLPSVSLDPLRLSIQKIDILTIKANTTLOITTCRGQDLD 60  
Db 1 MESKALLAVALWFCVETRAASVGLPGDFLHPKPLSTQKIDILTIANTTLOITTCRGQDLD 60

Qy 61 WLWPNNGSGSRQREVEVTEC--SDGLFCKTLTIPKVIQNDGTGAYKCFYRETDLASVIYVYV 118  
Db 61 WLWPNNAQDSERVLVTECGGSDSIFCKTLTIPRVGNDTGAYKCSYRDVDIASTVYVYV 120

Qy 119 QDYRSPFIASVDQGVVYITENKNTVTPICLGSISNLNVLSCARYPEKRFVDPGNRIS 178  
Db 121 RDRSPFIASVDQGVVYITENKNTVTPICRGSISNLNVLSCARYPEKRFVDPGNRIS 180

Qy 179 WDSKKGFTIPSYMLSYAGMVFCEAKINDESYQSTWYIVVYVYVYVYVYVYVYVYVYVYV 238  
Db 181 WDSIGFTIPSYMLSYAGMVFCEAKINDETYSYIVVYVYVYVYVYVYVYVYVYVYVYVYV 240

Qy 239 GEKVLNCTARTLNVLGIDFNWYPSGKHQHKLVNRDLTKTQSGSEMKKFLSTLITIDGVT 298  
Db 241 GEKVLNCTARTLNVLGIDFTWHPSPSKSHHKLVNRDVRFPFGTVAKMFLSTLITIESVT 300

Qy 299 RSDQGLVTCASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVIRIPAKYLGYP 358  
Db 301 KSDQGEVTCVASSGRMIKKNRTFVRVHTKPFIAFGSGMKSILVEATVGSQVRIPVKYLSP 360

Qy 359 PPEIKWYKNGIPLESNTIKAGHVLTIWESERTGNYTILTNPISEKQSHVSVYVYV 418  
Db 361 APDIKWYRNGRPIESNTIMVIGDELTIMETVTERDAGNYTILTNPISEKQSHVSVLVN 420

Qy 419 VPPQIGKSLISPVDSYQYGTQTLCTVVAIPPHPHIHWOLEECANEPQSAVSVN 478  
Db 421 VPPQIGKALISPVDSYQYGTQTLCTVVAIPPHPHIHWOLEECASVRPQ----TS 476

Qy 479 PYCEEWSRVEDFOGKNKIEVKNQFALIIEGKNKTVSTLVIQANVSALYKCEAVNKVR 538  
Db 477 PYACKERHVEDFOGKNKIEVTKNOYALIEGKNKTVSTLVIQANVSALYKCEAINKAGR 536

Qy 539 GERVISHTVRGPBITLOPMQPTQESVSLWCTADRTSTFENLTWYKGPQPIPIHVGEL 598  
Db 537 GERVISHTVRGPBITVQPAQPTQESVSLWCTADRTSTFENLTWYKGSQATSVHMGES 596

Qy 599 PTPVCKNLDLTKLNATMFSNSTDILIMELKNASLODQGDYVCLADRTKTKRHHCVVQ 658  
Db 597 LTPVCKNLDLTKLNGTWMFSNSTDILIVAFQNASLODQGDYVCSADKTKTKRHHCVVQ 656

Qy 659 LTVLERVAPITITGNLENQTTISIGESIEVSVCTASGNPPQIMWFKDNETLVEDSGIVLKDG 718  
Db 657 LTVLERNAPITITGNLENQTTIGETIEVTCASGNPPTPHITWFKDNETLVEDSGIVLRDG 716

Qy 719 NRNLTIIRVRKEBGLVTCQACSVLGCAYVEAFPIIEGAQEKTNLEIILVGTAVIAMFF 778  
Db 717 NRNLTIIRVRKEBGLVTCQACNVLCARAEFTPIIEGAQEKTNLEIILVGTAVIAMFF 776

Qy 779 WLLLVILRTVVRANGELKTYLSIYMDDELPLDEHCEPLPYDASKWEPFRDLKLGK 838  
Db 777 WLLLVILRTVVRANGELKTYLSIYMDDELPLDEHCEPLPYDASKWEPFRDLKLGK 836

Qy 839 PLGRGAFQVTEADAFGIDKATCTRTVAVKMLKEGATHSEHRLMSELKILIHIGHLVN 898  
Db 837 PLGRGAFQVTEADAFGIDKATCTRTVAVKMLKEGATHSEHRLMSELKILIHIGHLVN 896

Qy 899 VNLGACTKPGGLMWIVFCKFGLNLTSLYRSKNEFVYKTKGARFQKQKDYVGAIPVD 958  
Db 897 VNLGACTKPGGLMWIVFCKFGLNLTSLYRSKNEFVYKTKGARFQKQKDYVGAIPVD 956

Qy 959 LKRLDSITSSQSSASSGFFVEEKSLSVDEEEASEELYKDFLTLEHLICYSFQVAKGMF 1018

Db 957 LKRLDSITSSQSSASSGFFVEEKSLSVDEEEASEELYKDFLTLEHLICYSFQVAKGMF 1016

Qy 1019 LASRKCITHRDLAARNILLSEKNVVKICDFGLARDIYKDDPVYVRKGDARLPLKMMAPETIF 1078

Db 1017 LASRKCITHRDLAARNILLSEKNVVKICDFGLARDIYKDDPVYVRKGDARLPLKMMAPETIF 1076

Qy 1079 DRVTIISQDVMSCFGLLWEIFSLGASYPGVKIDEEFCRRLEKGTWRAPDYTTPEMYQT 1138

Db 1077 DRVTIISQDVMSCFGLLWEIFSLGASYPGVKIDEEFCRRLEKGTWRAPDYTTPEMYQT 1136

Qy 1139 MLDCWHGEPQRTFSELVEHLGNLQANQDQDKYIVLPISETLSNEEDSGLSLTSP 1198

Db 1137 MLDCWHGEPQRTFSELVEHLGNLQANQDQDKYIVLPISETLSNEEDSGLSLTSP 1196

Qy 1199 VSCHEEVECDPKPHYDNTAGISQYLONSKRKSRPVSVKTFEDIPLEPEVKVIPDDNQ 1258

Db 1197 VSCHEEVECDPKPHYDNTAGISQYLONSKRKSRPVSVKTFEDIPLEPEVKVIPDDNQ 1256

Qy 1259 DSGMVLASBEELKTLTDRTKLSPSGMVPSPKSRVASEGNSQTSYGQSVHSDDTDTTV 1318

Db 1257 DSGMVLASBEELKTLTDRTKLSPSGMVPSPKSRVASEGNSQTSYGQSVHSDDTDTTV 1316

Qy 1319 YSSEAEALLKLEIGVQTGSTAQILQDPSGTTL 1351

Db 1317 YSSEAEALLKLEIGVQTGSTAQILQDPSGTTL 1340

## RESULT 7

Q677M1 CHICK

ID Q677M1\_CHICK PRELIMINARY;

PRT; 1348 AA.

AC Q677M1; 2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Vascular endothelial growth factor receptor 2.

GN Name=Flk1;

OS Gallus gallus (Chicken)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus

OX NCBI\_TaxID=9031;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Hashimoto T.;

RT "Modulation of retinal neurogenesis by vascular endothelial growth

factor."

RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY382882; AAR26285.1; -; mRNA.

DR GO: GO:0016020; C:membrane; IEA.

DR GO: GO:0003779; F:actin binding; IEA.

DR GO: GO:0005524; F:ATP binding; IEA.

DR GO: GO:0004574; F:protein serine/threonine kinase activity; IEA.

DR GO: GO:0005021; F:vascular endothelial growth factor activity; IEA.

DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.

DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IEA.

DR InterPro; IPR001589; Actbind\_actnin.

DR InterPro; IPR003599; Ig.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003598; Ig\_c2.

DR InterPro; IPR000719; Prot\_kinase.

DR InterPro; IPR001824; RecepttyrknsIII.

DR InterPro; IPR002290; Ser\_thr\_kinase.

DR InterPro; IPR001245; Tyr\_kinase.

DR InterPro; IPR008266; Tyr\_kinase\_AS.

DR InterPro; IPR009134; VEGFR.

DR InterPro; IPR009136; VEGFR2.

DR Pfam; PF00047; ig; 1.

DR PRINTS; PR01832; VEGFRECSPTOR.

DR PRINTS; PR01834; VEGFRECSPTOR2.

DR ProDom; PD000001; Prot\_kinase; 2.

DR SMART; SM00409; IG; 7.

DR SMART; SM00408; IGC2; 5.

DR	SMART; SM00220; S_TKC; 1.	DR	831	LGRGAFQVTEADAFGIDKTATCRTVAVVWMLKEGATHSEHRLMSELKILIHGHHLNVV	890
DR	SMART; SM00219; TYKKG; 1.	Qy	900	NLLGACTKGGPLMWIVVECKGNLSTYLSRSENEVPYKTKGARFGKGD-VYGAIPVD	958
DR	PROSITE; PS00019; ACTININ_1; 1.	Db	891	NLLGACTKGGPLMWIVVECKGNLSTYLSRSENEVPYKTKGARFGKGD-VYGAIPVD	950
DR	PROSITE; PS00835; IG_LIKE_5.	Qy	959	LKRRLDSITSSOSSASSGFFVEEKSLSDVVEEEA-PEDLYKDFLTLEHLICYSFQVAKGME	1017
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.	Db	951	LKQRLDSITSSOSSASSGFFVEEKSLSDVVEEEA-PEDLYKDFLTLEHLICYSFQVAKGME	1010
DR	PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.	Qy	1018	FLASRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVYKRGDARLPLKWMAPETI	1077
DR	PROSITE; PS00109; PROTEIN_KINASE_TVR; 1.	Db	1011	FLASRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVYKRGDARLPLKWMAPETI	1070
KW	Receptor_TVR_KIN_III; 1.	Qy	1078	FORVYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRRLEKGRMAPDYTTPEMYQ	1137
Qy	SEQUENCE 1348 AA; 150121 MW; F6EP0F21CA839D6E CRC64;	Db	1071	FORVYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRRLEKGRMAPDYTTPEMYQ	1130
Query Match	71.8%; Score 5091; DB 2; Length 1348;	Qy	1138	TMLDCHWGHPSPPTSELSVEHLGNLLQANAQDGKDYIVLPISSETLSMEEDSGLSLPTS	1197
Best Local Similarity	72.1%; Pred. No. 2.3e-292;	Db	1131	TMLDCHWGHPKQRPTELSVEHLGNLLQANVRQDGKDYIVLPISVSLNMEEDSGLSLPTS	1190
Matches 981; Conservative 140; Mismatches 223; Indels 16; Gaps 13;		Qy	1198	PVSCMBEEBVCDPKPHYDNTAGISQYLQNSKRSRPSVSKTFEDIPLEBPEVKVIPPDDNQ	1257
Qy	1 MESKVLLAVALMCLVETRAASVGLPSVSLDPLRLSTOKDILTITKANTLLOITCRGORDLD	Db	1191	PASCKEEBEVCDPKPHYDNTAGISQYRQSGKRSRPSVSKTFEDIPLE-VTTVKVQEEHQ	1249
Db	1 MELGPLRVLVLCCLP--PVFTGL-FISMDOPTLSIQKSVLTITNDTLNITCSGORAVY	Qy	1258	TDGMYLASEELKLTLEDRTK-LSPSPGGMVPSKRSRESVASEGNSQTSYGSGVHSDDTDT	1316
Qy	61 WLWPNQSGEQRVEVTECSGDLFCFKTLTIKPKVIGNDTGAYKCFYRETDLASVIYVQVD	Db	1250	TDGMYLASEELKLTLEDRTK-LSPSPGGMVPSKRSRESVASEGNSQTSYGSGVHSDDTDT	1309
Db	58 WSNPNQSNAEKRLAVTGCSDGPFCKTLTLRLVIGNDTGDRCLYGDGSOAATIIYVQVD	Qy	1317	TVYSSEAEALLKLEIGVOTGSTAQILQDPSGTTLSPPV 1356	
Qy	121 YRSPFIASVSDQGVVITENKNTVVIKCLGSIINLVSLCARYPEKRFVDPGNRISWD	Db	1310	MVCSSEDTLLCAQEAAS-PFLPRVHGLVHDSAPLVSPPL 1348	
Db	118 YRSPFVTSVGDQLGIVIT--KNTVVPVCLGTVSNLVSLHAKYPEKVFVPDGSISWD	RESULT 8			
Qy	181 SKGFTIPSPMISYAGMVFCEAKINDESQSIIMYIVVVVGYRIYDVVLSPSHGIELSVGE	VGF2_COTJA			
Db	176 NKKGFTIPSHLSIYAGMVFCEAKINDESQSVIYVAVVGYRIYDLTMNPHYQVELAVGE	ID VGF2_COTJA STANDARD; PRT; 1348 AA.			
Qy	241 KLVNCTARTELNVGIDFNWEPSSKHQHKLVNRDLKTOSGSEMKEFLSTLITDGVTRS	AC P52583;			
Db	236 KLVNCTVRELNVGIDFNWEPSSKHQHKLVNRDLKTOSGSEMKEFLSTLITDGVTRS	DT 01-OCT-1996 (Rel. 34, Created)			
Qy	301 DOGLYTCASSGLMTKNSFTVRVHEKPFVAFSGMESLVEATVGRVIRIPAKYLYGPPP	DT 01-OCT-1996 (Rel. 34, Last sequence update)			
Db	293 DKGRYTCASSGRMNKNSYFIHESPFHL-EKMNENVEMKLGDTVSIPIVKFKGYSP	DT 13-SEP-2005 (Rel. 48, Last annotation update)			
Qy	361 EIKWYKNGIPLSENHITKAGHVTIMEVSRDTCNTVILTNPISEKQSHVSVLVVYVP	DE Vascular endothelial growth factor receptor 2 precursor (EC 2.7.1.112)			
Db	352 EAKWYKNGKVMANHTVKGTLVITDATEKAGNTVTVLTNPNQKQKRTHTLLVNVVP	GN Name-KDR; Synonym=EKL, FLK-1;			
Qy	421 PQIGKSLSPVDSYOGTTLCTVYAIPLPPHHILHWQLEEECANEPSSQAVSVTNPY	OS Coturnix coturnix japonica (Japanese quail).			
Db	412 PQIGENALWAPVDSYKSGTQTLCTIYAVPPPAALWQLEEECTFPQKRVLCANPY	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Qy	481 PCBEWRSVEDFOGKNIENKQNFALIEGKNTVSTLVIQAANVSALYKCEAVNKVGRGE	OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
Db	472 ACKRWKVISERKGNQVEI-KQVVVTIAGTKTSTLVIQAANVSALYRCMATNRAGSSE	OX NCBI_TaxID=93934;			
Qy	541 RVLSFHVTRGPETLQPDWQPTQESVSLWCTADRTFFENLWYKLGQPLPHVUGELPT	RN [1] NUCLEOTIDE SEQUENCE.			
Db	531 RVLSFHVTRGLEINLQPSQLTEKNTSLQCTADKFTFEKLSWYKLSAHALQTPFGGLPM	RP TISSUE=Embryo;			
Qy	601 PVCKNLDTLWKLNAWF-SNSTNDILIMELKNASLODQGDYVCLADRTKKEHCVVROL	RC MEDLINE=97017121; PubMed=8863722; DOI=10.1016/0378-1119(96)00159-X;			
Db	591 PVCKNLDALQKLNATVNTNGENVTLLEILNLSLQDGDYVCTADQKKAQTHCLVUKHL	RA Eichmann A., Marcelle C., Breant C., le Douarin N.M.;			
Qy	660 TVLERVAPITIGNLENTTIGSIEBVSCTASGNPPFQIMFWKDNLTLEDSDGIVLKGN	RT "Molecular cloning of Quak 1 and 2, two quail vascular endothelial			
Db	651 TVQEPNAPTLVGNLENTTIGTIEVSCVNGIPPPNITWPKNGETLFDSDGIVLKGN	growth factor (VEGF) receptor-like molecules.";			
Qy	720 RNLTIIRVRKEDBLVTCQACSVLGCACVAFPIIEGAQKNTLEIILVGTAVIAMFFW	RL Mech. Dev. 42:33-48(1993).			
Db	711 KTLTIIRVRKEDGGLYVCLACNLTLGCKKAEAYFSEGAEBEKTNLEIILVGTAVIAMFFW	RN [2] NUCLEOTIDE SEQUENCE OF 910-1348.			
Qy	780 LLLVILIRTVKANGELKTYLSIYMDPELPLDEHCERLPYDASKWFFPRDLKGRP	RP TISSUE=Spinal cord;			
Db	771 LLLVILIRTVKANGEMKTYLSIYMDPELPLDEHCERLPYDASKWFFPRDLKGRP	RC MEDLINE=93378866; PubMed=8396413; DOI=10.1016/0925-4773(93)90096-G;			
Qy	840 LGRGAFQVTEADAFGIDKTATCRTVAVVWMLKEGATHSEHRLMSELKILIHGHHLNVV	RA Eichmann A., Marcelle C., Breant C., le Douarin N.M.;			
		RT "Two molecules related to the VEGF receptor are expressed in early			
		endothelial cells during avian embryonic development.";			
		RL Mech. Dev. 42:33-48(1993).			
		RN [3] NUCLEOTIDE SEQUENCE OF 764-880, AND CHARACTERIZATION.			
		RP TISSUE=Embryo;			
		RC MEDLINE=95301109; PubMed=7781909; DOI=10.1006/dbio.1995.1180;			
		RA Flamme I., Breier G., Risau W.;			
		RT "Vascular endothelial growth factor (VEGF) and VEGF receptor 2 (flk-1)			
		are expressed during vasculogenesis and vascular differentiation in			

RT the quail embryo.";  
 RL Dev. Biol. 169:699-712(1995).  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE OF 1023-1079.  
 RX PubMed=1281306;  
 RA "Molecular cloning of a family of protein kinase genes expressed in  
 RT the avian embryo".  
 RL Oncogene 7:2479-2487(1992).  
 CC -1- FUNCTION: Receptor for VEGF or VEGFC. Has a tyrosine-protein  
 CC kinase activity. The VEGF-kinase ligand/receptor signaling system  
 CC plays a key role in vascular development and regulation of  
 CC vascular permeability.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + a protein  
 CC tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: In all endothelial tissues during onset of  
 CC vascularization. In later development, present in lung, heart,  
 CC intestine and skin.  
 CC -1- DEVELOPMENTAL STAGE: Expressed in whole mesoderm at onset of  
 CC gastrulation. From day 2, confined to endothelial tissues and  
 CC expression continues to be widespread throughout vascularization  
 CC until E9 where it becomes restricted to specific regions such as  
 CC the spinal chord and heart valves.  
 CC -1- INDUCTION: In vitro, it is induced by basic fibroblast growth  
 CC factor (bFGF), uniquely in the first 24 hours of cell culture.  
 CC -1- SIMILARITY: Belongs to the Tyr protein kinase family. CSF-1/PDGF  
 CC receptor subfamily.  
 CC -1- SIMILARITY: Contains 7 Ig-like C2-type (immunoglobulin-like)  
 CC domains.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 DR ENBL; X83288; CAA58268.1; -; mRNA.  
 DR ENBL; S65205; AAB28127.1; -; mRNA.  
 DR ENBL; S78345; AAB34594.1; -; mRNA.  
 DR PIR; JCA953; S51656.  
 DR HSP; P35968; IVR2.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig C2.  
 DR InterPro; IPR000719; Prot kinase.  
 DR InterPro; IPR001824; RecepttyrkinIII.  
 DR InterPro; IPR001245; Tyr pkinase.  
 DR InterPro; IPR008266; Tyr\_pkinase\_AS.  
 DR InterPro; IPR009134; VEGFR.  
 DR InterPro; IPR009136; VEGFR2.  
 DR Pfam; PF00047; Ig; 1.  
 DR PRINTS; PR01832; VEGFRECEPTR.  
 DR PRINTS; PR01834; VEGFRECEPTR2.  
 DR ProDom; PD000001; Prot kinase; 2.  
 DR SMART; SM00408; IGC2; 1.  
 DR SMART; SM00219; TyrKc; 1.  
 DR PROSITE; PS00835; IG LIKE; 5.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00011; PROTEIN KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE\_TYR; 1.  
 DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 KW Angiogenesis; ATP-binding; Developmental protein; Differentiation;  
 KW Glycoprotein; Immunoglobulin domain; Kinase; Nucleotide-binding;  
 KW Phosphorylation; Receptor; Repeat; Signal; Transmembrane;  
 KW Tyrosine-protein kinase.  
 FT SIGNAL 1 20 Potential.  
 FT CHAIN 21 1348 Vascular endothelial growth factor  
 FT receptor 2.  
 FT TOPO\_DOM 21 756 Extracellular (Potential).  
 FT TRANSMEM 757 776 Potential.  
 FT TOPO\_DOM 778 1348 Cytoplasmic (Potential).  
 FT DOMAIN 43 106 Ig-like C2-type 1.  
 FT DOMAIN 138 202 Ig-like C2-type 2.

FT	DOMAIN	220	312	Ig-like C2-type 3.
FT	DOMAIN	320	405	Ig-like C2-type 4.
FT	DOMAIN	412	534	Ig-like C2-type 5.
FT	DOMAIN	540	651	Ig-like C2-type 6.
FT	DOMAIN	658	744	Ig-like C2-type 7.
FT	DOMAIN	825	1155	Protein kinase.
FT	NP BIND	831	839	ATP (By similarity).
FT	ACT_SITE	1021	1021	By similarity.
FT	BINDING	859	859	ATP (By similarity).
FT	CARBOHYD	43	43	N-linked (GlcNAc). (Potential).
FT	CARBOHYD	47	47	N-linked (GlcNAc). (Potential).
FT	CARBOHYD	63	63	N-linked (GlcNAc). (Potential).
FT	CARBOHYD	93	93	N-linked (GlcNAc). (Potential).
FT	CARBOHYD	138	138	N-linked (GlcNAc). (Potential).
FT	CARBOHYD	153	153	N-linked (GlcNAc). (Potential).
FT	CARBOHYD	201	201	N-linked (GlcNAc). (Potential).
FT	CARBOHYD	240	240	N-linked (GlcNAc). (Potential).
FT	CARBOHYD	290	290	N-linked (GlcNAc). (Potential).
FT	CARBOHYD	310	310	N-linked (GlcNAc). (Potential).
FT	CARBOHYD	365	365	N-linked (GlcNAc). (Potential).
FT	CARBOHYD	386	386	N-linked (GlcNAc). (Potential).
FT	CARBOHYD	513	513	N-linked (GlcNAc). (Potential).
FT	CARBOHYD	556	556	N-linked (GlcNAc). (Potential).
FT	CARBOHYD	603	603	N-linked (GlcNAc). (Potential).
FT	CARBOHYD	613	613	N-linked (GlcNAc). (Potential).
FT	CARBOHYD	622	622	N-linked (GlcNAc). (Potential).
FT	CARBOHYD	666	666	N-linked (GlcNAc). (Potential).
FT	CARBOHYD	688	688	N-linked (GlcNAc). (Potential).
FT	CARBOHYD	710	710	N-linked (GlcNAc). (Potential).
FT	CONFLICT	865	865	A -> S (in Ref. 3).
FT	SEQUENCE	1348 AA;	150306 MW;	A5E4194A76PD5FB3 CRC64;

Query Match 70.9%; Score 5031.5; DB 1; Length 1348;  
 Best Local Similarity 72.1%; Pred. No. 7.7e-289;  
 Matches 960; Conservative 144; Mismatches 212; Indels 15; Gaps 11;

QY	1	MESKLLAVALMVCVETRAASVCLPSVSLDPLRLSLTQKDLITLTKANTLITLCRGQRLD	60
DB	1	MELGRLRLVTLVLLCLAPVAGL---FISMDQPTLSIQKSVLITTTNDLTNITSCGQAVY	57
QY	61	WLWPNNSQSGEQRVEVTEGSDGLFCKTLTIPIKVINNDTGAYKCYFRETDLASVYVYVD	120
DB	58	WSWPNNQSSVEKRLAVTGCSEGFCKTLTLRLVIGNDTGDRCLYGDSSQAATVYVYVD	117
QY	121	YRSPFIASVSDQHGVIYITENKNTVIVPCLSISNLSVCLARYPEKRFVPGNRIISWD	180
DB	118	YRSPFVTSVGDQLGIVIT--KNKTVVPCLVSNLNSLHAKYPEKVPVDPGKSIWD	175
QY	181	SKKGFTIPSYMISYAGMVFCEAKINDESYQSIMYVYVYGYRYDVVLSHSGIELSVGE	240
DB	176	NKKGFTIPSHLNYAGMVFCEAKINDESYQSVIYVAVVGYRYDVTMNPYQVELAVGE	235
QY	241	KLVLNCTARTELNVGIDFNWEPSSKHQKLVNRDLKTOSGSEMKKFLSTLTIDGVTRS	300
DB	236	KLVLNCTVTELNVGIDFRNDYFSEIKERRATI--RDLKTTAG-EIKTFVSTLIESVNL	292
QY	301	DQGLYTCAAASSGLMTKNSFTVRHKEKPFVAFSGMESLVEATVGERVRIPAKYLVGPPP	360
DB	293	DKGRYTCAAASSGRMNKNSYFIHESPFTHL-EKMNENVEMKLGDTVSPVFKGYPPP	351
QY	361	EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGNVTVILTNPISKEQSHVSVVYVP	420
DB	352	EAKWYKNGKVINAHHTVYKLGVALVITEATEKDAGNYTVVLTNTNKKQKHTFTLLVNP	411
QY	421	PQIGKSLISPVDSYQVGTQTTLCTVYALPPPHIHWYQLEEEECANEPASQVSVNTPY	480
DB	412	PQIGENALMAPVDSYKSGTQALTCTIYAVPPPAVLWYQLEEECTFSQKVLGNPY	471
QY	481	PCBEWRSVEDFOGKNKIEVNKNQFALIEGKNKTVSTLVIOAANVSALYKCEAVNKVGRGE	540
DB	472	ACKRWKVISERKGNQVEI-KQRVVVIAAGTKTVSTLVIOAANVSALYRCMATNRAGSSE	530
QY	541	RVISFHVTRGPEITLOPDMQPTQESVSLWCTADRSTFENLTWYKLGQPLPIHVGLPT	600



Db 531 RVISFHVTRGLEINLQPSQLTEKNTSLOCTADKFTFEKLSWYKLSHVSQTPFGGLPM 590  
Qy 601 PVCKNLDLTKLNATMFS-NSTNDILLMELKNLSLOQGDYVCLADRKTKKKBVVQRQL 659  
Db 591 PVCKNLDALQKLNATSVNGENVTLIELIRNLSLOQGDYVCLADKKAOTKCHLVKHL 650  
Qy 660 TVLERVAPTTIGNLENTTIGSIEVSTASGNPPQIMFWKDNETLVEDSGIVLKDGN 719  
Db 651 TVQEPHLRVLGNLENTTIGETIEVLCVNGVPPENITWFKNSETLFDSGIVLKDGN 710  
Qy 720 RNTIRVRKEDGLYTCQACSVLGCYAKVAPFPIIEGAQBKTNLEIILVGTAVIAMFFW 779  
Db 711 KTLTIRVRKEDGLYTCACNLGCKKAEAFVSQGAEBKTNLEIILVGTAVIAMFFW 770  
Qy 780 LLLVILITVRKANGELKTYGILSVMDPELDELDEHCERLPYDASKWEPFRRLKLGKP 839  
Db 771 LLLVILITVRKANGGMKTYGILSVMDPEVDEHCERLPYDASKWEPFRRLKLGKP 830  
Qy 840 LGRGAFQVTEADAFGIDKATCTRTAVAKMLKEGATHSEHRALMSELKILIHGHILNVV 899  
Db 831 LGRGAFQVLEADAFGIDKATCTRTAVAKMLKEGATHSEHRALMSELKILIHGHILNVV 890  
Qy 900 NLLGACTKPGGPMVIVFCKPGLNLTSLYLSKRNFPVYKTKGARFRQGD-KTVGAIPVD 958  
Db 891 NLLGACTKPGGPMVIVFCKPGLNLTSLYLSKRNFPVYKTKGARFRQGD-KTVGAIPVD 950  
Qy 959 LKRLDSITSSQSASSGFFVEEKSLSVDEEEA-PEDLYKDFLTLEHLICYSFOVAKGME 1017  
Db 951 LKRLDSITSSQSSTSSGFFVEEKSLSVDEEEDAGSDLCNPLTMEDLICYSFOVARGME 1010  
Qy 1018 FLASRKCIHRDLAARNILLSEKNVWIKIDFGLARDYKDPDYVRKGDARLPLKWMAPETI 1077  
Db 1011 FLASRKCIHRDLAARNILLSDNNVVKIDFGLARDYKDPDYVRKGDARLPLKWMAPETI 1070  
Qy 1078 FDRVYTIQSDVMSFGVLLWEIFSLGASPYPGVKIDEEFCRRLEKGTMRAPDYTTPEMYQ 1137  
Db 1071 FDRVYTIQSDVMSFGVLLWEIFSLGASPYPGVKIDEEFCRRLEKGTMRAPDYTTPEMYQ 1130  
Qy 1138 TMLDCHWGEPSPRTSELVEHLGNLLOANQDGKDYIVLPISFELSMEEDSGLSPTS 1197  
Db 1131 TMLDCHWGDPKQRPTESELVEHLGNLLOANQDGKDYIVLPISFELSMEEDSGLSPTS 1190  
Qy 1198 PVSCEEEECVDPKFHYDNTAGISQYLSQNSKRSPVSVKTFPEDIPLERPEVKVIPDDNQ 1257  
Db 1191 PASCKEEECVDPKFHYDNTAGISQYLSQNSKRSPVSVKTFPEDIPL-VTTVKVVQSEMQ 1249  
Qy 1258 TDSGMVLASELKTLEDRTK-LSPSPGGMVPSKRSVSASEGNSQTSQYSGYHSDTDT 1316  
Db 1250 TDSGMVLASELKTLEDRTK-LSPSPGGMVPSKRSVSASEGNSQTSQYSGYHSDTDT 1309  
Qy 1317 TVYSSEAEELL 1327  
Db 1310 MVCSSDTELL 1320

RESULT 9  
QSGIT4\_BRARE PRELIMINARY; PRT; 1357 AA.  
AC QSGIT4;  
DT 10-MAY-2005 (TrEMBLrel. 30, Created)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)  
DE Receptor tyrosine kinase Kdr.  
GN Names=kdr; Synonyms=kdr;  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Habeck H., Langhoff J., Vogel A.M., Trowe T., Koblizek T.I.,

RA Schulte-Merker S.;  
RT "Synergistic signaling of vegf receptors is required for  
RT vasculogenesis in zebrafish";  
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY523399; AA522270.1; -; mRNA.  
DR Ensembl; ENSDARG0000017321; Danio rerio.  
DR ZFIN; ZDB-GENE-041001-112; kdrb.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0005021; F:vascular endothelial growth factor receptor. .; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; IEA.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003598; Ig\_c2.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR001824; RecepttyrkinsIII.  
DR InterPro; IPR002290; Ser thr\_kinase.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR InterPro; IPR008266; Tyr\_kinase\_AS.  
DR InterPro; IPR009134; VEGFR.  
DR InterPro; IPR009137; VEGFR3.  
DR Pfam; PF00047; ig; 2.  
DR PRINTS; PR01832; VEGFRECEPTOR.  
DR PRINTS; PR01835; VEGFRECEPTR3.  
DR ProDom; PD000001; Prot\_kinase; 2.  
DR SMART; SM00409; IG; 7.  
DR SMART; SM00220; S\_TKc; 1.  
DR SMART; SM00219; Tyrc; 1.  
DR PROSITE; PS00835; IG\_LIKE; 5.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
KW Kinase; Receptor.  
SQ SEQUENCE 1357 AA; 152952 MW; 16A477FBF08F3B33 CRC64;  
Query Match 46.0%; Score 3264.5; DB 2; Length 1357;  
Best Local Similarity 50.3%; Pred. No. 3.7e-184;  
Matches 678; Conservative 195; Mismatches 370; Indels 105; Gaps 26;  
Qy 30 DLPRLSIQKDLITKANTTLOITCQRDLMLPNQSGSORVEVTCSS-DGLFCKTL 88  
Db 30 DPTLITKIKINASDLOITCQRGRILEWSTPHNRTLSSTLTLTIDCSGDLFCSTL 89  
Qy 89 TIPKVIQNDTGAYKCFY----RETDLASV-IYVYQDYRSPFTASVSDQGVVYITENKN 143  
Db 90 TLSKAVANETGEYRCFYKSLPKEDGKTSVAVYVFIQDYTFV-RIAQDYDVVFIREGEQ 148  
Qy 144 KTVVIPCGLGISNLNVSLCARYPEKRFVDPGNRISWDSKKGFTTIPSYMTISAGMVFCEAK 203  
Db 149 --VWIPCLVSVEDLVNTLYTKYPVKELSTDGKEVWDSRRGFILPSRVVSAGVVCQT 206  
Qy 204 INDESQSTWYIVVVGVRIVDVVLSPLSHGIELSVCKLAVNCTARTELNVGDENWEYP 263  
Db 207 IRNETFQSSPYIVAVVVGKYLIDLTLSPOHE-RLTVGERLILNCTAHTELNVGIDFQMT 265  
Qy 264 SSKHQHKKLVNRDLKT---QSGSEMKK-----FLSTLTIDGVTRSDQGLYTCAASSGLM 314  
Db 266 ----HEKRSVNGSMSTSRVKTSSNKKLWNSLESLNTLVENLTNDTGEYICTASSGQM 321  
Qy 315 TKKNSTFVRVHEKPFVAFSGMESLVEATVGE-RVRIPAKYLGYPPEPEIKWYKNG--IPL 371  
Db 322 QKIAQASLVYKEKPFALSDQLWQTVKAGDAEAKILVKYAYPEPAVRVKNQDLIVL 381  
Qy 372 ESNHTIK--AGHVLTIMEVSEDTGNTVILNPLSKESKOSHVSLSLVVYVPPQIGKSLI 429  
Db 382 RDEYRMKFRVGHLLTYIGVTEKDAGNTVMTNKIKTEEQRTFTQLVNDLPRIFEKQVS 441  
Qy 430 SPVDSYQYGTOTLTCTVVAIPPPHHIHWYQLEBECANEPQAV--SVTNPYPCBEWRS 487



```
Db 442 LDRDVHMGSPITLTCTASGSSPVTIKWMPREDC---PVRLPKSDTRMAKCKWRE 498
Qy 488 VEDPQGNKIEVKNQKOPALIEGKNKTSTLVIQAANVSALYKCEAVNKGGRGVISFHV 547
Db 499 MSNNTGKNPL---ISQTSVDETLTKTSLKIQKAVDHALYRCIATNKGQDQVIVSQV 555
Qy 548 TRGPEITLQDMQPTQESVSLCTADRSFENLTWYKLG-----PQPLPIHVGEPLTP 601
Db 556 TRFLNLSVSSPIEGQDVMRCVADRLLYNLRVYRVANVANVHDPAPPCDTL--- 612
Qy 602 VCKNLDTLWKLNAFMFS-NSTNDILIMELKNASLDQDGVVCLAQDKTKKRHCWVRQLT 660
Db 613 ---TLSHLHPNVTVSGLOGTNVTLDMPINATMMDQGLYACQVEIVGTNEKTCLLHNR 669
Qy 661 VLVERAPITGNLENQTTSGESIEVSTAGSNPPQIMFKONETLVEDSGIVLKDGNR 720
Db 670 LRALEMSRIVTNTLDQRVNVSDSTLVCESGTPPTVTWKDQNTMEGSGVILKRSNR 729
Qy 721 NLTIIRVRKEDEGLYTQACSVLGCACVFAFFIEGAEKTNLEIILVGTAVIAMPFWL 780
Db 730 VLTQIRVKEDSGLYICTACNQCCSESEARI SVDGAEKXNVNELIMPICAVIAMPFWL 789
Qy 781 LVLVILRTVRANGELKTVLTVMDPDELPLDEHCELRPYDASKWEPDRDLKLGKPL 840
Db 790 LTVFVIRNRKPNDDGDKTGLTSLSDDDMPDEHCELRITYDASKWEPDRDLKLGKPL 849
Qy 841 GRGAFQVIBADAFGIDKATCTRTAVVAKMLKEGATHSEHRALMSELKILHIGHLNVN 900
Db 850 GRGAFQGVVEATVAGIEKATCTTAVVAKMLKEGATSEYHRALMSELKILHIGHLNVN 909
Qy 901 LLGACTKPGGLMIVVEFCFKNLSTYLRSKRNEFVYKTKGARFQGGY-----VG 953
Db 910 LLGACTKQGGFLMIVVEYCKHGNLSYLKSGREYSYKRTTPRPNRKEVQODEDPREG 969
Qy 954 AIPDLKRLDSTSSQSSASSGFVEEKSLSVDEEERAPEDLYKDFLTLEHLICYFOVA 1013
Db 970 DLGLGTSTRIDICTGTAVCTRTGEGTYTKTLQD---EQESSD---WDHLTMDLISYFQVA 1024
Qy 1014 KGMEFLASRCKIHRDLAARNILSEKNVVKICDFGLARDIYKDPYVVRKGDARLPLKWA 1073
Db 1025 KGMEFLASRCKIHRDLAARNILSENSVVKICDFGLARDIYKDPYVVRKGDARLPLKWA 1084
Qy 1074 PETIFDRVYITQSDVMSFGVLLMEIFSLGASPPGVKIDEEFCRRLKEGTRMRAPDYTP 1133
Db 1085 PETIFDRVYITQSDVMSFGVLLMEIFSLGASPPGVKIDEEFCRRLKEGTRMRAPDYTP 1144
Qy 1134 EMYQTMDCWHGEPSPQPTSELSVEHLGNLLQANAQDQKDYIVLPISETLSMEEDSGLS 1193
Db 1145 EYVQTMDCWLDRLDRPTTQLVEHLGNLLQASAQDQKDYIPLTNGE----- 1193
Qy 1194 LPTSPVSCMEEEVVC-----DPKFHYDNT--AGISQYLONSKRKSRPVS 1235
Db 1194 -----MEELVAPHLNVTSKRSPYAGNTEAQLHYDNAPPFGFPQOMNSS---GVFVN 1242
Qy 1236 VKTFEDIPLEPEYKVIIPDNDQSDGMVLAHEELKLTLEDRTKLSPSGMGVPSKRESVA 1295
Db 1243 MTGFVDIPLHTTW-----DGHVDCGVGLSREQMKALDRQAQRPLNFSPLLRCKSKESLA 1298
Qy 1296 SEGSNQTSQYQSGVHSDDTDTTVYSSEE 1323
Db 1299 SEGSNQTSQYQSGVHSDDAEAPIYANEE 1326
```

## RESULT 10

```
Q4RKW3_TETNG PRELIMINARY; PRT: 1328 AA.
ID Q4RKW3
AC Q4RKW3
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 1 SCAF15025, whole genome shotgun sequence.
DE (fragment).
```

```
GN ORFNames=GSTENG00032761001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Athouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Bernardis V.,
RA Craud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chappier C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nuebaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RL the early vertebrate proto-karyotype.";
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAE01015025; CAG10969.1; -; Genomic_DNA.
FT NON_TER 1328 1328
SQ SEQUENCE 1328 AA; 148349 MW; C7C3874FBAC62C1C CRC64;
```

Query Match 43.1%; Score 3058.5; DB 2; Length 1328;  
Best Local Similarity 47.0%; Pred. No. 5.7e-172;  
Matches 638; Conservative 203; Mismatches 400; Indels 117; Gaps 24;

```
Qy 48 TLQITCRGQRLDNLWPNQSGSEQREVEVTECS-DGLFCKTLTPKVIQNDTGAYCFYR 106
Db 6 SLELTCSGRQLHWTTPRTSS---RFSIGDCSGSLFCCTTLAISNATVNETQYQCSYK 61
Qy 107 ETDL-----ASVIVYVQDYRSPFIASVSDQHQGVVITENKNTVIVPCLGSIINLVSL 161
Db 62 DLKVEDGKTAATYVVEDYKVPFV-PLKKEHEVVFIRE--GEWVIPCRESSVDLNVTL 118
Qy 162 CARYPEKRFVDPGNRISWDSKKGFTIPSYMISVAGMVFCEAKINDESYQSIMIVVVVGY 221
Db 119 HTKYPNKLYPDGKTEWDAKGFVMPFSLISVAGVVPQQTQIGNETFKSPLIVAVVGY 178
Qy 222 RIYDVVLSPSHGTEL SVGKGLVNLCTARTELNVGIDFNMEY-----PSSKHQH 269
Db 179 KIYDLTLTPAK-TQLAAGERLVLSCTAITELVNGIEFNWTHSGOALASVNSSRPTVATPH 237
Qy 270 KKLNVRLDKTQSGEMKFLSTLITDGVTRSDOGLYTCASSGLMTKQNSTFVRVHEKPF 329
Db 238 KKKLWNSLSELS-----NTLTVENVTVDSHGEYTCSSASSGLMKEKSESVPVKYKPF 288
Qy 330 VAFSGMESLVEATGGERV-RIPAKYLGYPPEIKMYKNGIPLSNHTIKAG-HVLTIME 387
Db 289 IEIKPEWAKVLEVNVDQLTTPVTKYSAYPEPGYKWLKNGQPLKEDYRFQKLDNLVIRG 348
Qy 388 VSRDGTGNTVILTNPISKEQSHVSVLVVYVPPQIGKSLISFVDSYQYGTQTTLCTV 447
Db 349 VTETDAGLTVTLTKITKEEQRSLLQLVNVPPIIEKEVAADTVYPVSSPTLRCTA 408
Qy 448 YAIPPPPIHHWYQLEEEECANE-PSQAVSVTNPYPCSEWRSVSDFOGKNKIEVKNQFAL 506
Db 409 RGFPIPSHIEWOMWAKEDCPEAFQSRPNPNTLLQCSGWRISGNSGTHRAIEGVKKYCRF 468
Qy 507 IEGKNTVSTLVLOANVSLKCEAVNKVGRGRVISHVTRGPETITLQPDQPTQES 566
```





```
Db 191 PEP---VDGSVVWNNKKGWSIPRHIIQNTSTFGF-YCSISVQNSQHTSSIIYVQVIGL 246
Qy 222 RIYDVVLSPSHG-IELSVGKLVNCTARTELNVGIDFNWEYPSXKHQKLVNRLDKTQ 280
Db 247 KFEYFKLPEDSPVELMQGSLVNLCTALVDFTGVDFQWDYFCKK--ENRLASLQPLRN 304
Qy 281 SGSEMKFLSLTIDGVTGRSDQGLYTCAASGLMTKKNSTFVRVHEKPFVAFSGMESLV 340
Db 305 VLDEATEISSILSRINHLSDSGYTCWANTLEMKRELTTWIVVHEKPFISLDYRNGSVI 364
Qy 341 EATVGER-VRIPAKYLGYPPPEIKWKYKNGKPLPSNHTIK---AGHVLTIMEVSEDTGNY 396
Db 365 EAKEGQSVLSVKVSAIPSEIQYKNGKLISSKNSRFKVOQHSIQIRIDCKVQAGEY 424
Qy 397 TVILTN-PISEKQSHVSVVSVVPPQIGESKLISPVDSYQYGTOTLTCTVVAIPPHH 455
Db 425 MVLKNSPAALEKRLN-FTLIVNVPPQIHEKEAAPTNLYGKTRQILICTAGD-SPAS 482
Qy 456 IHMYQLEECANPEPQAV-----SVTNPPYCEWRSVEDFQGNKIEVKNKQFALIEG 509
Db 483 ISQWRPSPCDLERTRALRRRGRGRDQSPF-CHNWMDLDPERHAVNPIESIDTLTQWVG 541
Qy 510 KNTVSTLVTOANVSALYKCEAVNKVGRGERSVISHVTRGPE---ITLQPDNQPTQES 566
Db 542 KEKTVGRWVIONASVPAMYKCLAEKVKGDERLIYFYVTYIPGFDIEMEPSEDPLEQDL 601
Qy 567 VSLWCTADRTFENLTWYKLPQPLPIHVCELPVCKNLD---TLWKLNATMFSNSTND 623
Db 602 VQJKNADNFTYENLRWYLDPTQVP-----PELDCKSLHQTATFEGQLSFQTTISNWW 655
Qy 624 ILIMELKNASLODQGVYVCLAQDRKTKRCHVVRQLTVLERVAPTITGNLENQTTIGRS 683
Db 656 VLQLNTINIQLQDEGNTVCVQNRRTGVKCHRKYIPVKAMEAPRYRHNPNTNHTVNVSES 715
Qy 684 IEVSCTAGSNPPQIMWFKDNETLVEDSGIVLQGRHNLTIKVRKDEGLYTCQACSVL 743
Db 716 LQWNCDEGVEFPFQSLQWFKDNQHLQISGLLODSNRTLQIQRVEDAGLYTCSACNQK 775
Qy 744 GCAKVEAFFIIEGAQKTNLEIILVGTAVIAMFFWLLVILRTVKRANGELKTYLS 803
Db 776 GCVQSATSVSVIGSDKTNVEIILGTGVIATFFWVLLVIFCNVKNVPADIKTYLS 835
Qy 804 INVMDPELDEHCERLPYDASKWEPPDRDLKGLGPGAFQGVIEADAFGIDKATCR 863
Db 836 IIMDPGEVPLEQCEVLPYDSSQWEISRDRLRGLGVGHGAFKVIETASIFGHDKSSAN 895
Qy 864 TVAVKMLKEGATSEHRLMSELKILIHGHHLNVNVLGACTKPGGLMVIVFECFKN 923
Db 896 TVAVKMLKEGATSEHRLMSELKILIHGHHLNVNVLGACTKPGGLMVIVFECFKN 955
Qy 924 LSTYLSKRNEFPVYKTKGARFQGDYVGAI PVDLKRRLDSTTSQSSASSGFVEEKS 983
Db 956 LSNFLRAKREFFLPYRDRSPKTS-----QVRMEAGQASQSE-----HQPS 999
Qy 984 SDVEEEAP-EDLYKDFLTLEHLICYSFQVAKMEFLASRKCCHRDLAARNILLSKNVV 1042
Db 1000 SSTNPRVTVDLWKPLTIEDLICYSFQVARGMEFLASRKCCHRDLAARNILLSNNVV 1059
Qy 1043 KICDFGLARDIYKDPYVRKGDARLPLKWMAPTIFDRVYTTQSDVWSFGVLLWEIPLG 1102
Db 1060 KICDFGLARDIYKDPYVRKGNARLPLKWMAPESIFDKVYTSQSDVWSFGVLLWEIPLG 1119
Qy 1103 ASPYGVKIDEBFCRLKEGTRMRAPDYTPPEWQTMDCWHCEPSQRPTEFSLVHLGN 1162
Db 1120 ASYPGIGIDEDFCRLKQKOTRMRAPDNASPEIYIGIMLACWQGEPRPTFPALVILGD 1179
Qy 1163 LLOANAQODQKOYIVLPISETLSMEEDSGLSLPTSPVSCMEEVECD--PKHYDNT--A 1218
Db 1180 LLOENSLPE-----IPENVQSSEDGDFSQASSRPPSQEIRLACNTLPTRYNCVPPA 1233
Qy 1219 GISQYLONSKRKRPVSVKTFEDIPLBEPEVKVIPDNQTDGSMVLASELKTLEDTK- 1277
Db 1234 GCMVWGPSSTCHSR---VKTFEELPEMTSHKT-QHDSQTDGSMVLASDELEFHEKHRG 1289
```

---

```
Qy 1278 --LSPSFGQWPSK--SRESVASEGS-----NOTSG---YOSGY-HSDDTDTTVY 1319
Db 1290 AMLTATTGOSTDRLLSCFSVSSSGGGLLRPVFTQLSGQTFYNNYGHLSSEGVSDY 1349
Qy 1320 --SSEA 1324
Db 1350 FSSDQA 1356

RESULT 13
Q5MD89 BRARE PRELIMINARY; PRT; 1357 AA.
AC Q5MD89;
DT 01-FEB-2005 (T-EMBLrel. 29, Created)
DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (T-EMBLrel. 29, Last annotation update)
DE Flt4.
GN Names=flt4;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lawson N., Moore J.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY833404; AAV93318.1; --; mRNA.
DR ZFIN; ZDB-GENE-980526-326; flt4.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0005021; F:vascular endothelial growth factor receptor; IEA.
DR GO; GO:0004668; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . ; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001824; RecepttyrkinasIII.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR008266; Tyr_pkinase.
DR InterPro; IPR009134; VEGFR.
DR InterPro; IPR009137; VEGFR3.
DR Pfam; PF00447; ig; 2.
DR PRINTS; PR01832; VEGFRRECEPTOR.
DR PRINTS; PR01835; VEGFRRECEPTR3.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00409; IG; 6.
DR SMART; SM00408; IGC2; 6.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00835; IG LIKE; 5.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
SQ SEQUENCE 1357 AA; 153177 MW; D68C197828BE278E CRC64;

Query Match 40.2%; Score 2855.5; DB 2; Length 1357;
Best Local Similarity 44.9%; Pred. No. 6.2e-160;
Matches 623; Conservative 222; Mismatches 415; Indels 127; Gaps 37;

Qy 22 VGLP-----SVSLDPLRLSTQKDLITLKANTTQITCGQRDLWLWP----- 64
Db 13 IGIPFFSGLVGFSMSPPTLDNTKQDLVINANDTLNITCGQRILDWSPEESLSKVEFT 72
Qy 65 ---NNQSGE---QRVEVTECSDEL---FCKTLTPKVIQNDTGAYKCFYRET----- 108
Db 73 DRQGGQSPTDTPGYREIRLKEC-QGVAGPKYCKLILITNAQANDSGYRCFYNDIKAVID 131
```

109 -DLASVIYVQDYRSFPIASVSDHGVVITENKNTVWIPCLGISNLNVL--CARY 165  
132 GTTHAASFVVRDPEHFIRKGNMDMETIFITDSETH-IEVPCLVSDPDKLVLFLSLVPY 190  
166 PEKRFVDPGNRIISWDSKGGFTIPSYMI-----SYAGMVFCEAKINDESIQSIYVWVGY 221  
191 PEP---VDGSVVWNNKKGWSIPRHIIQNTSTFIFG-YCSISVQNSQHTSSIVVQVIGL 246  
222 RIYDVVLSPHG-IELSVGEKVLNCTARTLNVGIDFNWEYPSKQHKHKLNVRLDKTQ 280  
247 KPYEFKLPEDSPVELMQGESVLNCTALVDNTGVDFQWDYDGGK--ENRLASLQPLRN 304  
281 SGSEMKKFLSTLTIDGVTSDOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFGSGMESLV 340  
305 VLDEATISILSTRNTHLDDSGYTCWANTLEMKRELTTVVVHEKPFISLDRNGSVI 364  
341 EATVGER-VRIAPKYLGPPEIKWYKNGIPLESNHTIK--AGHVLTIMEVSESDTGNV 396  
365 EAKGQKSVLNVKVSAYSPSEIQWYKNGKLISKNSSRFKQVQHSIQIRDVCKQDAGEY 424  
397 TVILTN-PISKEKSHVSVLVVYPPQIYGEKSLISPVDSYQYGTOTLTCTVVAIPPHH 455  
425 MLVLKNSPAALKEKLN-FTLVVNPQIHEKEAAPPTNLKYGKTRQLTCTADG-SPPAS 482  
456 IHWYQLEEBECANPESQAV-----SVTNVPYCEEWSVEDFCQGNKIEVKNQFALIEG 509  
483 ISQWRPWSPPCDLERTRALRRGRDQSPF-CHNMWDLDPHEAVNPISIDITLQWVDG 541  
510 KNTKTVLTVQAANVSALYKCEAVNKVGRSERVIFSHVTRGPE---ITLQPDMDPTQES 566  
542 KEKTVGRVVLQNASVPAMYKLAENRVGKDERLIYFVTTIPGDFDIEMEPSDPLEQDL 601  
567 VSLMCTADRSTFENLWYKGPQLPIHVGELPPVCKNLD---TLWKLNAVFSNSTND 623  
602 VOLKCNADNFYENLWYRLDPQTVP-----PELDCSLHQVATFLEGQLSFTQTSNNW 655  
624 ILMELNKSILQDGDVYCLAQDKTKRHCVRVQLTVLVRVAPTITGNLENQTTSGES 683  
656 VLQNLITNIQLQDGNVCEVQNNRTGVKCHRYIIPVKAMEAPRYRHNPTNHTNVNSES 715  
684 IEVSTAGNPPPOIMWFKNETLVEDSGVLKDGNNLTIRVRKEDEGLYTCQACSVL 743  
716 LQWNCVDEGTFPQLSNFKNQLPHQISLLODSNRTLSIQVRVEDAGLYTCSACNQK 775  
744 GCAKVEAFFIEGAQEKTNLEIILVGTAVIAMFFWLLVILRTVVRANGELKTYLS 803  
776 GCQSSATSVIGSDDKTNVEIILGTGVIAIFFVLLVILFVNCVKNVPADIKTYLS 835  
804 IVMDDPELDEHCERLIPYDAKWEFPRDRILKGLKPLGRGAFQVIEADAFGDKTATCR 863  
836 IIMDPGEVPLEEQEYLPYDSSQWEISRDRRLRLGKVLGHGAFGVIEASTFGHDKKSAN 895  
864 TVAVKMLKEGATSEHAKMSLKIILHGHNVNLLGACTKPGGFLVIVIEFCFKGN 923  
896 TVAVKMLKEGATSEHAKMSLKIILHGHNVNLLGACTKPGGFLVIVIEYCKGN 955  
924 LSTYLRKRNFEVYKTKGARFRQKDYVGAIPVDLKRRLDSTTSQSSASSGPFVEKSL 983  
956 LSNFLRAKREFFLPYDRSPKTSQ-----QVRMTIEAGQASQE-----HQPST 999  
984 SDVEEBEAP-EDLYKDFLTLEHLICYSFOVAKMGFLASKRCHRDLAARNILLSEKVV 1042  
1000 SSTNPPRTVTDLWKTLETTIEDLTCYSFQVARGMEFLASKRCHRDLAARNILLSENNV 1059  
1043 KICDFGLARDYKDPDYVRKGDARLPLKWMAPETIPRVVYTIQSDVMSFGVLLWEIFSLG 1102  
1060 KICDFGLARDYKDPDYVRKGNARLPLKWMAPESIFDKVYTSQSDVMSFGVLLWEIFSLG 1119  
1103 ASPYGVKIDBEFCRRLEKGTREAPDYTTPEMTQTMDCWHGSPSPRFTFSELVEHLGN 1162  
1120 ASPYGGIQIDDFCKRLKDGTRMEAPDNASPEIIVGIMLACWQGEPRPRPTFPALVEILGD 1179  
1163 LLQANAQDGKDYIVLPISETLSMEEDSGLSLPTSPVSCMEEEVECD--PKFHYDNT--A 1218

1180 LLQENSLPE-----IPFNVQSSEDDGFQASRRPSSQSEIRLACNTLTRYNCVPPFA 1233  
1219 GISQYQNSKRKRPSVSTFEDIPLEBPVKVVPDQNTDQSGMVLASELKLTDRTK- 1277  
1234 GCVMVGPSSSTCHSR--VKTFFELPMEMTSKHT-QHDSQTDQSGMVLASDELPERFEKHG 1289  
1278 --LSPFGMGVPSK--SRESVASEGS-----NOTSG---YQSGY-HSDDTDITVY 1319  
1290 AMUTTATTGOSTDRLLISCPVSSVSSGGGLLRPVFFTLQSGQTFYNNYHGLSEGVSDY 1349  
1320 --SSEEA 1324  
1350 FSSDQA 1356

RESULT 14  
VGPR3\_MOUSE  
ID VGPR3\_MOUSE STANDARD; PRT; 1363 AA.  
AC P35917;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Vascular endothelial growth factor receptor 3 precursor (EC 2.7.1.112)  
DE (VEGFR-3) (Tyrosine-protein kinase receptor FLT4).  
GN Name=Flt4; Synonyms=Flt-4;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muroidae; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J;  
RX MEDLINE=93330572; PubMed=8393164;  
RA Finnerty H., Kelleher K., Morris G.E., Bean K., Merberg D.M., Kriz R.,  
RA Morris J.C., Sookdeo H., Turner K.J., Wood C.R.;  
RT "Molecular cloning of murine FLT and FLT4.";  
RL Oncogene 8:2293-2298(1993).  
RN [2]  
RP NUCLEOTIDE SEQUENCE OF 1033-1072.  
RX PubMed=1319394;  
RA Galland P., Karamyheva A., Mattei M.-G., Rosnet O., Marchetto S.,  
RA Birnbaum D.;  
RT "Chromosomal localization of FLT4, a novel receptor-type tyrosine  
RT kinase gene.";  
RL Genomics 13:475-478(1992).  
CC -I- FUNCTION: Receptor for VEGFC. Has a tyrosine-protein kinase  
CC activity.  
CC -I- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + a protein  
CC tyrosine phosphate.  
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -I- TISSUE SPECIFICITY: Expressed in adult lung and liver, and in  
CC fetal liver, brain, intestine and placenta.  
CC -I- SIMILARITY: Belongs to the Tyr protein kinase family. CSF-1/PDGF  
CC receptor subfamily.  
CC -I- SIMILARITY: Contains 7 Ig-like C2-type (immunoglobulin-like)  
CC domains.  
CC  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC  
CC EMBL; L07296; AAA40077.1; -; mRNA.  
CC PRP; I58375; I58375.  
CC HSPR; P35968; IVR2.  
CC Ensembl; ENSMUSG0000020357; Mus musculus.  
CC MGI; MGI:95561; Flt4.  
CC GO; GO:0005615; C:extracellular space; TAS.  
CC GO; GO:0016021; C:integral to membrane; TAS.  
CC GO; GO:0005515; F:protein binding; IPI.





**This Page Blank (uspto)**



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2006, 18:19:09 ; Search time 146.073 Seconds  
(without alignments)  
4078.762 Million cell updates/sec

Title: US-10-633-742-6

Perfect score: 7095

Sequence: 1 MESKVLVALWLCVETRAA.....GSTAQLQPSGTTLSPPV 1356

Scoring table: BLOSUM62

Gapop 10.0., Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*  
9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7095	100.0	1356	4	AAY97783 Human KDR
2	7092	99.9	1356	4	AAB62475 Human VEG
3	7092	99.9	1356	4	AAY97576 Human Flk
4	7092	99.9	1356	6	ABR40196 Human vas
5	7092	99.9	1356	7	ADD08954 Human VEG
6	7092	99.9	1356	7	ADM79007 Human VEG
7	7092	99.9	1356	7	ADF45097 Human kin
8	7092	99.9	1356	8	ADH17130 Human vas
9	7092	99.9	1356	8	ADQ39806 Human myo
10	7092	99.9	1356	8	ADQ39808 Human myo
11	7092	99.9	1356	8	ADR46648 Human as
12	7092	99.9	1356	8	ADT92353 Human vas
13	7092	99.9	1356	9	ADV90285 Protease-
14	7092	99.9	1356	9	ADY59385 Human VEG
15	7092	99.9	1356	9	ADZ26561 Human VEG
16	7091	99.9	1356	5	AAU79426 Human Kin
17	7091	99.9	1356	8	ADG70543 Human KDR
18	7090	99.9	1356	7	ABU64302 Human KDR
19	7087	99.9	1356	2	AAW80997 Human rec
20	7087	99.9	1356	5	AAU79427 Human Kin
21	7087	99.9	1356	5	AAU79430 Human Kin
22	7087	99.9	1356	5	AAU79429 Human Kin
23	7070	99.6	1356	2	AAR26999 Novel typ
24	7070	99.6	1356	2	AAW59275 Human KDR

25	7060	99.5	1354	9	ADZ00446	Adz00446 VEGFR-2.
26	6806	95.9	1306	8	ADQ39807	Adq39807 Human myo
27	6183.5	87.2	1343	8	ADR32338	Adr32338 Rat optim
28	6143.5	86.6	1343	8	ADR32351	Adr32351 Rat recep
29	6127.5	86.4	1345	7	ABM79009	Abm79009 Murine Fl
30	6106.5	86.1	1367	6	ABR40197	AbR40197 Murine va
31	6097.5	85.9	1367	2	AAR31377	Aar31377 Murine flk
32	6097.5	85.9	1367	2	AAR37504	Aar37504 Murine fl
33	6097.5	85.9	1367	2	AAR44996	Aar44996 Murine fl
34	6097.5	85.9	1367	2	AAR67537	Aar67537 Mouse flk
35	6097.5	85.9	1367	2	AAR97420	Aar97420 Murine fl
36	6097.5	85.9	1367	2	AAW19875	Aaw19875 Murine fl
37	6097.5	85.9	1367	2	ADP90725	Adp90725 Mouse foe
38	6097.5	85.9	1367	2	AAU08618	Aay08618 Murine fl
39	6097.5	85.9	1367	5	ABG70917	Abg70917 Mouse rec
40	6097.5	85.9	1367	5	AAE25820	Aae25820 Murine re
41	6097.5	85.9	1367	9	ADM87791	Adm87791 Murine fe
42	6097.5	85.9	1367	9	AEBA6056	Aeb46056 Murine fl
43	6094.5	85.9	1367	2	AAR28041	Aar28041 flk-1.3/
44	6091.5	85.9	1367	2	AAR67817	Aar67817 Flk1 rece
45	6066.5	85.5	1367	2	AAR54046	Aar54046 Sequence

#### ALIGNMENTS

RESULT 1  
AAY97783  
ID AAY97783 standard; protein; 1356 AA.

AC AAY97783;

DT 22-AUG-2001 (first entry)

DE Human KDR/Flk-1 protein.

KW Immunoglobulin-like domain; Ig-like domain; retinal neovascularisation;  
KW vascular endothelial growth factor receptor; VEGF receptor; inflammation;  
KW psoriasis; rheumatoid arthritis; haemangioma; leiomyoma; angiofibroma;  
KW diabetic retinopathy; endometriosis; macular degeneration; cancer;  
KW dimerisation inhibitor; therapy; KDR/Flk-1; human.

OS Homo sapiens.

XX WO200142284-A2.

PN 14-JUN-2001.

PD 07-DEC-2000; 2000WO-GB004693.

XX 07-DEC-1999; 99GB-00028950.

PA (METR-) METRIS THERAPEUTICS LTD.

XX Pappa H;

PI WPI; 2001-381631/40.

XX Novel proteins that prevent dimerization of vascular endothelial growth  
XX factor receptors and for treating diseases in which the receptor plays a  
XX role e.g. inflammation, cancer, diabetic retinopathy, psoriasis.

PT Claim 11; Fig 3; 83pp; English.

XX This sequence represents the human KDR/Flk-1 protein. The invention  
XX relates to a protein comprising the amino acid sequence of the fourth  
XX immunoglobulin (Ig)-like domain of a vascular endothelial growth factor  
XX (VEGF) receptor or a variant of the protein that retains the ability to  
XX bind to a VEGF receptor. The protein, its functional equivalent and DNA  
XX encoding it are useful for treating a disorder whose pathology is  
XX dependent upon a VEGF family-mediated pathway, including inflammation,  
XX psoriasis, rheumatoid arthritis, haemangioma, leiomyoma, diabetic  
XX retinopathy, angiofibrosis, endometriosis, macular degeneration, retinal



CC growth factor receptor-2 (VEGFR-2) receptor protein and a neuropilin-1  
CC (NP-1) receptor protein. One method comprises introducing a sample  
CC comprising the compound to the receptor protein and allowing the compound  
CC to bind to the complex. Signaling through VEGFR-2 is enhanced in the  
CC presence of the NP-1 co-receptor. The methods of the invention can be  
CC used for identifying novel pro- and anti-angiogenic compounds. The  
CC present sequence represents the human VEGFR-2 receptor protein  
XX  
XX Sequence 1356 AA;

Query Match 99.9%; Score 7092; DB 4; Length 1356;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESKVLAVLWLCVETRAASVGLPSVSLDPLRLSIQKIDILTIKANTLIQITCRGQDLD 60  
DB 1 MQSKVLAVLWLCVETRAASVGLPSVSLDPLRLSIQKIDILTIKANTLIQITCRGQDLD 60

QY 61 WLMPNNSGSEORVEVEVTECDGLFCKLTIPKVIIGNDTGAYKCFYRETDLASVIYVQD 120  
DB 61 WLMPNNSGSEORVEVEVTECDGLFCKLTIPKVIIGNDTGAYKCFYRETDLASVIYVQD 120

QY 121 YRSPFIASVDQHGVIYITENKNTVVIICLGSISNLNLSLCARYPEKRFVPGNRI SWD 180  
DB 121 YRSPFIASVDQHGVIYITENKNTVVIICLGSISNLNLSLCARYPEKRFVPGNRI SWD 180

QY 181 SKGFTTIPSYMI SVAGMVFCEAKINDESYOSIMYIVVVVGYRIYDVVLSPSHGIELSVGE 240  
DB 181 SKGFTTIPSYMI SVAGMVFCEAKINDESYOSIMYIVVVVGYRIYDVVLSPSHGIELSVGE 240

QY 241 KLVLNCTARTALNVGIDFNVEYSSKHOKKLVNRDLKTQSGSEMCKFLSTLTIDGVTRS 300  
DB 241 KLVLNCTARTALNVGIDFNVEYSSKHOKKLVNRDLKTQSGSEMCKFLSTLTIDGVTRS 300

QY 301 DOGLYTCAASGLMTKKNSTFVRVHEKPFVAFSGMESLVEATGERVRIPAKYLYGPPP 360  
DB 301 DOGLYTCAASGLMTKKNSTFVRVHEKPFVAFSGMESLVEATGERVRIPAKYLYGPPP 360

QY 361 EIKWYKNGIIPLESNHTIKAGHVLTIMEVSRDGTGNTVILTNPISEKQSHVSLVYVP 420  
DB 361 EIKWYKNGIIPLESNHTIKAGHVLTIMEVSRDGTGNTVILTNPISEKQSHVSLVYVP 420

QY 421 PQIGKSLIPVDSYQGTTLCTCTVYAIIPPHIHWYQLEECANESQAVSTNYP 480  
DB 421 PQIGKSLIPVDSYQGTTLCTCTVYAIIPPHIHWYQLEECANESQAVSTNYP 480

QY 481 PCBEWRVDFQGGNKLEVKNQFALLIEGKNKTVSTLVIQANVSALYKCEANVKVGRGE 540  
DB 481 PCBEWRVDFQGGNKLEVKNQFALLIEGKNKTVSTLVIQANVSALYKCEANVKVGRGE 540

QY 541 RVISFHVTRGPEITLQDPMQTEQSVSLMCTADRTFENLTWYKLGQPLPIHVGLPT 600  
DB 541 RVISFHVTRGPEITLQDPMQTEQSVSLMCTADRTFENLTWYKLGQPLPIHVGLPT 600

QY 601 PVCKNLDLWKLNATMFNSNTDILIMELKNASLODQDGVCLADRTKTKRHCVRQLT 660  
DB 601 PVCKNLDLWKLNATMFNSNTDILIMELKNASLODQDGVCLADRTKTKRHCVRQLT 660

QY 661 VLBERVAPITGNLENQTTISGESIEVSTAGSNPPPOIMWFKONETLVEDSGIVLKDGNR 720  
DB 661 VLBERVAPITGNLENQTTISGESIEVSTAGSNPPPOIMWFKONETLVEDSGIVLKDGNR 720

QY 721 NLTIIRRVKEDEGLYTQCAQSVLGCACVFAFFIIEGAQEKTNLEIILVGTAVIAMFFWL 780  
DB 721 NLTIIRRVKEDEGLYTQCAQSVLGCACVFAFFIIEGAQEKTNLEIILVGTAVIAMFFWL 780

QY 781 LLVLIILTVKANGELKTYLSIVMDPDELPLDEHCERLPYDASKWEFPRDRILKGLPL 840  
DB 781 LLVLIILTVKANGELKTYLSIVMDPDELPLDEHCERLPYDASKWEFPRDRILKGLPL 840

QY 841 GRGAFGVIEADAFGIDKTATCTRTAVKMLKEGTHSEHRLMSELKILIHGHILNVN 900  
DB 841 GRGAFGVIEADAFGIDKTATCTRTAVKMLKEGTHSEHRLMSELKILIHGHILNVN 900

QY 901 LLGACTKPGGPLMVIIVEFCCKEGLNSTYLRSKRNEFVYKTKGARFRQGDYVGAIPVDLK 960  
DB 901 LLGACTKPGGPLMVIIVEFCCKEGLNSTYLRSKRNEFVYKTKGARFRQGDYVGAIPVDLK 960

QY 961 RRLDSITSSOSSASSGQFVEEKSLSDVVEEERAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
DB 961 RRLDSITSSOSSASSGQFVEEKSLSDVVEEERAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020

QY 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKPDYVYRKGARLPLKMAPETIFDR 1080  
DB 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKPDYVYRKGARLPLKMAPETIFDR 1080

QY 1081 VYTIQSDVWSFGVLLWEIFSLGASPYGVKIDEEFCERLKEGTRMRAPDYTTPEMYQTM 1140  
DB 1081 VYTIQSDVWSFGVLLWEIFSLGASPYGVKIDEEFCERLKEGTRMRAPDYTTPEMYQTM 1140

QY 1141 DCWHGSPSRPTSELSVEHLGNLLQANAQDQGDYIVLPISETLSMEEDSGLSLPTSPVS 1200  
DB 1141 DCWHGSPSRPTSELSVEHLGNLLQANAQDQGDYIVLPISETLSMEEDSGLSLPTSPVS 1200

QY 1201 CMEEBVCDDPKFYDNTAGISQYQLNSKRKSRPVSVKTFEDIPLERPEVKVIPPDDNQDTS 1260  
DB 1201 CMEEBVCDDPKFYDNTAGISQYQLNSKRKSRPVSVKTFEDIPLERPEVKVIPPDDNQDTS 1260

QY 1261 GMVLASEELKLTEDRTKLSFGMVPSPKSRRESVASEGSGNOTSGYQSGYHSDDDTDTTVYS 1320  
DB 1261 GMVLASEELKLTEDRTKLSFGMVPSPKSRRESVASEGSGNOTSGYQSGYHSDDDTDTTVYS 1320

QY 1321 SEBAELKLIEIGVQGTGTAQILQPDGTTLSPPV 1356  
DB 1321 SEBAELKLIEIGVQGTGTAQILQPDGTTLSPPV 1356

RESULT 3  
AAY97576  
ID AAY97576 standard; protein; 1356 AA.  
XX AAY97576;  
AC AAY97576;  
DT 05-APR-2001 (first entry)  
XX Human Flk-1 transmembrane and extracellular domains.  
DE Human; angiogenic protein; wound healing; vascular tissue repair;  
XX peripheral arterial disease; critical limb ischemia; coronary disease;  
KW angiogenesis; tumour; inflammation; diabetic retinopathy; psoriasis;  
KW rheumatoid arthritis; autoimmune disease; allergy; cancer; therapy;  
KW infectious disease; neurodegeneration; Flk-1.  
XX Homo sapiens.  
OS WO200075163-A1.  
XX 14-DEC-2000.  
PD 01-JUN-2000; 2000MO-US014925.  
XX 03-JUN-1999; 99US-0137796P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA Rosen CA, Ruben SM, Hu J, Cao L;  
XX WPI; 2001-071057/08.  
XX New nucleic acid encoding angiogenic proteins, useful e.g. for promoting  
PT healing of wounds and treating peripheral arterial disease, critical limb  
PT ischemia or coronary disease.  
XX Example 48; Page 236-240; 244pp; English.  
XX This sequence is a human Flk-1 fragment, and was used in the isolation of  
CC

an angiogenic protein of the invention. The angiogenic proteins and the DNA sequences encoding them, are used to prevent, treat or ameliorate disease and to detect diseases, or susceptibility, by detecting mutations or the presence or amount of angiogenic protein expression. Particularly they are used to stimulate wound healing, growth of damaged bone and tissue, and for repair of vascular tissue, especially peripheral arterial disease, critical limb ischaemia or coronary disease. Antagonists of the sequences are used to inhibit angiogenesis in tumours and to treat inflammation (where associated with increased vascular permeability), diabetic retinopathy, rheumatoid arthritis or psoriasis. Agonists are also useful for stimulating (lymph)angiogenesis. The proteins are also used to identify specific binding agents (potential therapeutic agents) and to raise antibodies. The antibodies are useful as therapeutic (ant)agonists; for detection, purification and targeting of proteins for in vivo or in vitro diagnosis (including imaging) or for therapy (including when linked to e.g. a label or cytotoxin); and for immunotyping of cells, e.g. for detecting minimal residual disease or haematopoietic progenitor/stem cells. It is also contemplated that the sequences might be useful for treating a very wide range of other disorders, e.g. autoimmune diseases; allergy; cancer; infectious diseases (viral, bacterial, fungal or parasitic); neurodegeneration, also as chemotactic agents or for stimulating regeneration of the nervous system etc

Query Match 99.9%; Score 7092; DB 4; Length 1356;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1356; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

CC	1	MESKVLAVLWLCVETRAASVGLPSVSLDPLRLSIQKQILTIKANTLIQTICRGQDRLD	60
CC	1	MQSKVLAVLWLCVETRAASVGLPSVSLDPLRLSIQKQILTIKANTLIQTICRGQDRLD	60
CC	61	WLPNNGSGSEQRVEVTECDGLFCKTLTIPKVGNDTGAYKCFYRETDLASVIVYVQD	120
CC	61	WLPNNGSGSEQRVEVTECDGLFCKTLTIPKVGNDTGAYKCFYRETDLASVIVYVQD	120
CC	121	YRSPFTASVSDQHGVIYITENKNTVIPCGLGISNLNLSLCARYPEKRPVPGNRI	180
CC	121	YRSPFTASVSDQHGVIYITENKNTVIPCGLGISNLNLSLCARYPEKRPVPGNRI	180
CC	181	SKKGFTIPSMISYAGMVFCEAKINDESYSIMYIVVVGYRIYDVVLSPSHGIELSVGE	240
CC	181	SKKGFTIPSMISYAGMVFCEAKINDESYSIMYIVVVGYRIYDVVLSPSHGIELSVGE	240
CC	241	KLVLNCTARTELVNGIDFNWEYSPSSKHQKLVNRDLKTQSGSEMKKFLSTLTIDGVT	300
CC	241	KLVLNCTARTELVNGIDFNWEYSPSSKHQKLVNRDLKTQSGSEMKKFLSTLTIDGVT	300
CC	301	DQGLYTCASSGLMTKKNSTFVRHKEPFAFGSGMESLVEATVGERVRIPAKYLGVPPP	360
CC	301	DQGLYTCASSGLMTKKNSTFVRHKEPFAFGSGMESLVEATVGERVRIPAKYLGVPPP	360
CC	361	EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTNTVILTNPISEKQSHVSVLVVYVP	420
CC	361	EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTNTVILTNPISEKQSHVSVLVVYVP	420
CC	421	PQIGKSLISPVDYSQYGTQTTLCTVYALPPPHHHWVQLBEECANBPQAVSVTNPY	480
CC	421	PQIGKSLISPVDYSQYGTQTTLCTVYALPPPHHHWVQLBEECANBPQAVSVTNPY	480
CC	481	PCSEWSEVDFQGNKIEVKNQFALIEGKNTVSTLVIOANVSALYKCEAVNKVGRGE	540
CC	481	PCSEWSEVDFQGNKIEVKNQFALIEGKNTVSTLVIOANVSALYKCEAVNKVGRGE	540
CC	541	RVISFHVTRGPEITLQPDQPTQESVSLWCTADRTFFENLTWYKLGQPLPHVGE	600
CC	541	RVISFHVTRGPEITLQPDQPTQESVSLWCTADRTFFENLTWYKLGQPLPHVGE	600
CC	601	PVCNKLDTLWKLNATWFSNSTNDILIMELKNASLQDQGVCLAQDRKTKRHC	660
CC	601	PVCNKLDTLWKLNATWFSNSTNDILIMELKNASLQDQGVCLAQDRKTKRHC	660

QY	661	VLERVAPTITGNLENOTTISIGSEIYVSCASGNPPPOIMWFKDNETLVEDSGIVLKDGNR	720
DB	661	VLERVAPTITGNLENOTTISIGSEIYVSCASGNPPPOIMWFKDNETLVEDSGIVLKDGNR	720
QY	721	NLTIRVRKDEGLYTCQACSVLGCACVRAFFIIEGAQEKTNLEIILVGTAVIAMFFWL	780
DB	721	NLTIRVRKDEGLYTCQACSVLGCACVRAFFIIEGAQEKTNLEIILVGTAVIAMFFWL	780
QY	781	LLVILITVRKANGELKTYLSIVMDPDELPLDEHCERLPYDASKWEPPDRDLKLGKPL	840
DB	781	LLVILITVRKANGELKTYLSIVMDPDELPLDEHCERLPYDASKWEPPDRDLKLGKPL	840
QY	841	GRGAFQVTEADAFGIDKTATCTAVKMLKEGATHSEHRLMSELKILIHIGHLNVN	900
DB	841	GRGAFQVTEADAFGIDKTATCTAVKMLKEGATHSEHRLMSELKILIHIGHLNVN	900
QY	901	LLGACTKPGGLMVIIVFCKFGLNLSYLSKRNEFVYPTKGFARFQGDYVGAIPVDLK	960
DB	901	LLGACTKPGGLMVIIVFCKFGLNLSYLSKRNEFVYPTKGFARFQGDYVGAIPVDLK	960
QY	961	RLDLSITSSQSSASSGFEKSLSDVEEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA	1020
DB	961	RLDLSITSSQSSASSGFEKSLSDVEEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA	1020
QY	1021	SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR	1080
DB	1021	SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR	1080
QY	1081	VYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRRLKEGTRMRAPDYTTPEMTQML	1140
DB	1081	VYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRRLKEGTRMRAPDYTTPEMTQML	1140
QY	1141	DCWHGEPSPRPFSELVEHLGNLLQNAQDGKDYIVLPISETLSNEEDSGLSPTSPVS	1200
DB	1141	DCWHGEPSPRPFSELVEHLGNLLQNAQDGKDYIVLPISETLSNEEDSGLSPTSPVS	1200
QY	1201	CMEEEVCDPKFHYDNTAGISQYVLSNKRKSRPVSVKTFEDIPLEPEVKVIPPDDNOTDS	1260
DB	1201	CMEEEVCDPKFHYDNTAGISQYVLSNKRKSRPVSVKTFEDIPLEPEVKVIPPDDNOTDS	1260
QY	1261	GMVLASEELKTLIEDRTKLSFSGMVPKSRSEVASEGNSQTSQYSGYHSDDTDTTVYS	1320
DB	1261	GMVLASEELKTLIEDRTKLSFSGMVPKSRSEVASEGNSQTSQYSGYHSDDTDTTVYS	1320
QY	1321	SEAEELKLIETGVQGTGTAQILQPDGTTLSPPV 1356	
DB	1321	SEAEELKLIETGVQGTGTAQILQPDGTTLSPPV 1356	
RESULT 4			
ABR40196	ID	ABR40196 standard; protein; 1356 AA.	
XX	XX	ABR40196;	
XX	XX	AC	
XX	XX	01-AUG-2003 (first entry)	
XX	XX	Human vascular endothelial growth factor receptor-2.	
DE	DE	Human; vascular endothelial growth factor receptor-2; cytostatic;	
XX	KW	angiogenic; antiangiogenic; antiarthritic; antirheumatic; antisense;	
KW	KW	VEGFR-2; hyperproliferative disorder; cancer; rheumatoid arthritis;	
XX	KW	angiogenesis.	
XX	OS	Homo sapiens.	
XX	XX	WO2003029266-A1.	
PN	XX	10-APR-2003.	
PD	XX	26-SEP-2002; 2002WO-US030734.	
PP	XX		
XX	XX		

PR	28-SEP-2001; 2001US-00967655.
XX	(ISIS-) ISIS PHARM INC.
PA	Bennett CF, Watt AT;
XX	WPI; 2003-371980/35.
XX	N-PSDB; ACC71713.
DR	New compounds, particularly antisense oligonucleotides targeted to a
XX	nucleic acid encoding vascular endothelial growth factor receptor-2
PT	(VEGFR-2), useful for treating a disease/condition associated with VEGFR-
PT	2, e.g. cancer.
XX	
PS	Example 13; Page 90-97; 127pp; English.
PS	The present invention relates to novel antisense oligonucleotides
CC	(ACC71728-ACC71750 and ACC80101-ACC80155) targeted to Vascular
CC	Endothelial Growth Factor Receptor-2 (VEGFR-2) nucleotide sequence, and
CC	which inhibit the expression of VEGFR-2. The oligonucleotides are useful
CC	in compositions for treating a disease or condition associated with VEGFR
CC	-2, such as hyperproliferative disorder, e.g. cancer, a disease or
CC	condition involving angiogenesis, or rheumatoid arthritis. The present
CC	sequence is human VEGFR-2
XX	
SQ	Sequence 1356 AA;
	Query Match 99.9%; Score 7092; DB 6; Length 1356;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY	1 MESKVLALVALMCLVETRAASVGLSPSLDLPRLSIQKDILTIKANTTLOITCRGQRDL 60
DB	1 MQSKVLLAVALMCLVETRAASVGLSPSLDLPRLSIQKDILTIKANTTLOITCRGQRDL 60
QY	61 WLPNNQSGSEORVEVTECSDDLTKTLTPKVIGNDTGAYKCFYRETDLASVIYVVQD 120
DB	61 WLPNNQSGSEORVEVTECSDDLTKTLTPKVIGNDTGAYKCFYRETDLASVIYVVQD 120
QY	121 YRSPFTASVDQHGVVYTENKNKTWTIPLCGSISLNLSVCARYPEKRFPVDCNRI 180
DB	121 YRSPFTASVDQHGVVYTENKNKTWTIPLCGSISLNLSVCARYPEKRFPVDCNRI 180
QY	181 SKKGFTIPSYMTISYAGWFCEAKINDESYQSIMYIVVVVGYYRIYDVVLSHSHGIELSVGE 240
DB	181 SKKGFTIPSYMTISYAGWFCEAKINDESYQSIMYIVVVVGYYRIYDVVLSHSHGIELSVGE 240
QY	241 KLVLNCTARTELNVGIDFNWEYPSSKHQHKLVNRDLTKQSGEMKKFLSTLTIDGVTRS 300
DB	241 KLVLNCTARTELNVGIDFNWEYPSSKHQHKLVNRDLTKQSGEMKKFLSTLTIDGVTRS 300
QY	301 DOGLYTCASSGLMTKNSTFVRVHEKPVPFAGSGMESLVEATVGERVRIPAKYLGYPPP 360
DB	301 DOGLYTCASSGLMTKNSTFVRVHEKPVPFAGSGMESLVEATVGERVRIPAKYLGYPPP 360
QY	361 EIKWKNGKIPLESNTIKAGHVLTITMEVSERDTGNVTYILTNPISKQSHWSLVVYVP 420
DB	361 EIKWKNGKIPLESNTIKAGHVLTITMEVSERDTGNVTYILTNPISKQSHWSLVVYVP 420
QY	421 PQIGEKSLISPVDYSYGTTQTTLCTTVYAIPPHHHIHWYOLEBECANEPQAVSVTNPY 480
DB	421 PQIGEKSLISPVDYSYGTTQTTLCTTVYAIPPHHHIHWYOLEBECANEPQAVSVTNPY 480
QY	481 PCEWRVSDFOGGNKIEVNKNQFALIEGKNKTSTVLVIAANYALYKCAEVNKGVE 540
DB	481 PCBEWRVSDFOGGNKIEVNKNQFALIEGKNKTSTVLVIAANYALYKCAEVNKGVE 540
QY	541 RVISFHVTGPEITLPQDMQPTQEBSVSLWCCTADSTFENLTWYKLGPOLPHVHGELPT 600
DB	541 RVISFHVTGPEITLPQDMQPTQEBSVSLWCCTADSTFENLTWYKLGPOLPHVHGELPT 600
QY	601 PVCKNLDTLWLKNATMFNSTNDILIMELKNASIQDQGDYVCLAQDRKTKRCCHCVVRLT 660
DB	601 PVCKNLDTLWLKNATMFNSTNDILIMELKNASIQDQGDYVCLAQDRKTKRCCHCVVRLT 660



AC ABM79007;  
 XX  
 AC  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 XX  
 DE Human VEGF receptor-2 (KDR).  
 XX  
 KW Human; vascular endothelial growth factor receptor-2; VEGFR-2; KDR;  
 KW receptor; vaccine; genetic immunisation; gene therapy; cytostatic.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO2003073995-A2.  
 XX  
 XX PD 12-SEP-2003.  
 XX  
 XX PF 28-FEB-2003; 2003WO-US006256.  
 XX  
 XX PR 02-MAR-2002; 2002US-00090183.  
 XX  
 XX PA (SCRI ) SCRIPPS RES INST.  
 XX  
 XX PI Reisfeld RA, Niethammer AG, Xiang R;  
 XX  
 XX DR WPI: 2003-756753/71.  
 DR N-PSDB; ACF80601.  
 XX  
 XX  
 XX New DNA vaccine for eliciting an immune response against proliferating  
 PT endothelial cells comprising a DNA construct operably encoding a VEGF  
 PT receptor protein in a carrier, useful for inhibiting tumor growth or  
 PT angiogenesis.  
 XX  
 XX Claim 3; Fig 2; 56pp; English.  
 PS  
 XX The present sequence is the protein sequence of human vascular  
 CC endothelial growth factor receptor-2 (VEGFR2 or KDR). A claimed DNA  
 CC vaccine effective for eliciting an immune response against proliferating  
 CC endothelial cells comprises a DNA construct operably encoding a VEGF  
 CC receptor protein (e.g. KDR) in a carrier. The DNA construct may be a  
 CC naked DNA construct or incorporated into a plasmid vector or into an  
 CC attenuated bacterial vector such as attenuated *Salmonella typhimurium*.  
 CC The DNA vaccine is used in claimed methods of inhibiting endothelial cell  
 CC proliferation, inhibiting angiogenesis, and inhibiting tumour growth  
 XX  
 XX Sequence 1356 AA;  
 SQ  
 Query Match 99.9%; Score 7092; DB 7; Length 1356;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 1 MESKVLAVLWLCVETRAASVGLPSVSLDLPRLSIQKDLTIKANTTLQITCRGQRDL 60  
 1 MQSKVLLAVLWLCVETRAASVGLPSVSLDLPRLSIQKDLTIKANTTLQITCRGQRDL 60  
 61 WLWPNNGSGEQRVEVTECSGLPCKLTTPKVTGNDTGAYKCFYRETDLASVIYVVD 120  
 61 WLWPNNGSGEQRVEVTECSGLPCKLTTPKVTGNDTGAYKCFYRETDLASVIYVVD 120  
 121 YRSPFIASVDQHGCVVITENKNTVVIICLGSISNLNLSVCARYPEKRFVPGNRLSWD 180  
 121 YRSPFIASVDQHGCVVITENKNTVVIICLGSISNLNLSVCARYPEKRFVPGNRLSWD 180  
 181 SKGFTTIPSYMISYAGMVFCEAKINDESYQIMIVVVVGYRIYDVVLSPSHGIELSVGE 240  
 181 SKGFTTIPSYMISYAGMVFCEAKINDESYQIMIVVVVGYRIYDVVLSPSHGIELSVGE 240  
 241 KLVNCTARTELNVDGIDFNWEYPSGKHQKLVNRLDKLTQSGSEMKKFLSLTIDGVT 300  
 241 KLVNCTARTELNVDGIDFNWEYPSGKHQKLVNRLDKLTQSGSEMKKFLSLTIDGVT 300  
 301 DOGLYTCAASGLMTKKNSTFVRVHEKPFVAFGSGMESLVEATVGERVIRPAKLYGPP 360  
 301 DOGLYTCAASGLMTKKNSTFVRVHEKPFVAFGSGMESLVEATVGERVIRPAKLYGPP 360

RESULT 7

ADF45097

ID ADF45097 standard; protein; 1356 AA.



XX ADF45097;  
 XX AC  
 XX DT 12-FEB-2004 (first entry)  
 XX DE Human kinase VEGFR2.  
 XX KW Human; protein kinase; enzyme; inhibitor; VEGFR2.  
 XX OS Homo sapiens.  
 XX PN W02003081210-A2.  
 XX PD 02-OCT-2003.  
 XX PP 20-MAR-2003; 2003WO-US008725.  
 XX PR 21-MAR-2002; 2002US-0366892P.  
 XX PA (SUNE-) SUNESIS PHARM INC.  
 XX PI Prescott JC, Braisted A;  
 XX DR WPI; 2003-865136/80.  
 XX CC Identifying ligand binding to inactive conformation of target protein  
 CC kinase (T) comprises contacting the conformation modified (T) which  
 CC contains reactive group at binding site, with ligands and detecting  
 CC kinase-ligand conjugate formation.  
 XX CC Disclosure; SEQ ID NO 66; 260pp; English.  
 XX PS  
 XX CC The present invention relates to a method for identifying a ligand (L),  
 CC which binds to an inactive conformation of target protein kinase (T). The  
 CC method involves contacting inactive conformation of (T), which contains  
 CC or is modified to contain a reactive group at or near a binding site of  
 CC interest, with one or more ligand candidates capable of covalently  
 CC bonding to the reactive group thus forming a kinase-(L) conjugate (C).  
 CC The method is useful for identifying protein kinase inhibitors that  
 CC preferentially bind to inactive conformation of a target protein kinase.  
 CC The present sequence is a protein kinase which may be modified via an  
 CC amino acid substitution, for use in the method of the invention.  
 XX CC  
 XX SQ Sequence 1356 AA;  
 Query Match 99.98; Score 7092; DB 7; Length 1356;  
 Best Local Similarity 99.94; Pred. No. 0;  
 Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MESKVLAVALLVCVETRAASVGLPSVSLDLPRLSIQKDLITIKANTTLQITCRGORDLD 60  
 DB 1 MQSKVLAVALLVCVETRAASVGLPSVSLDLPRLSIQKDLITIKANTTLQITCRGORDLD 60  
 QY 61 WLPNNGSQSEQRVEVTECDGLFCKTLTIPIKVGNDTGAYKCFYRETDLASVIYVYVQD 120  
 DB 61 WLPNNGSQSEQRVEVTECDGLFCKTLTIPIKVGNDTGAYKCFYRETDLASVIYVYVQD 120  
 QY 121 YRSPFTASVSDQHGVIYITENKNTVVIICLGSISNLNLSLCARYPEKRFVPGNRIISWD 180  
 DB 121 YRSPFTASVSDQHGVIYITENKNTVVIICLGSISNLNLSLCARYPEKRFVPGNRIISWD 180  
 QY 181 SKKGFTIPSWIISVAGMVFCEAKINDBSYQIMYIVVVGRIYDVVLSPSHGIELSVGE 240  
 DB 181 SKKGFTIPSWIISVAGMVFCEAKINDBSYQIMYIVVVGRIYDVVLSPSHGIELSVGE 240  
 QY 241 KLVNCTARTELNVGIDFNWYEPSSKHQKLVNRDLKTQSGSEMKKFLSTLIDGVTWS 300  
 DB 241 KLVNCTARTELNVGIDFNWYEPSSKHQKLVNRDLKTQSGSEMKKFLSTLIDGVTWS 300  
 QY 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIPAKYLGYPPP 360  
 DB 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIPAKYLGYPPP 360

RESULT 8  
 ADH17130  
 ID ADH17130 standard; protein; 1356 AA.  
 XX

QY 361 EIKWYKNGIPLESNHTTIKAGHVLTIMEVSRDGTGNTVILTNPISEKQSHVSVLVVYVP 420  
 DB 361 EIKWYKNGIPLESNHTTIKAGHVLTIMEVSRDGTGNTVILTNPISEKQSHVSVLVVYVP 420  
 QY 421 PQIGKSLISPVDYSYQYGTGTTQTLCTVYAIPPPHHIHWYQLEBEECANESQAVSVTNPY 480  
 DB 421 PQIGKSLISPVDYSYQYGTGTTQTLCTVYAIPPPHHIHWYQLEBEECANESQAVSVTNPY 480  
 QY 481 PCSEWESVEDFOGKNKIEVNKNQFALIEGKNKTVSTLVIOAANVSALYKCEANVKVGRGE 540  
 DB 481 PCSEWESVEDFOGKNKIEVNKNQFALIEGKNKTVSTLVIOAANVSALYKCEANVKVGRGE 540  
 QY 541 RVISFHVTRGPEITLQPDMPQTEQESVSLWCTADRSSTFENLTWYKLGPOPLPIHVGELPT 600  
 DB 541 RVISFHVTRGPEITLQPDMPQTEQESVSLWCTADRSSTFENLTWYKLGPOPLPIHVGELPT 600  
 QY 601 PVEKNLDTLWKNATWFSNSTNDILIMELKNASLOQDGYVCLAQDRKTKKRHCVVRLQT 660  
 DB 601 PVEKNLDTLWKNATWFSNSTNDILIMELKNASLOQDGYVCLAQDRKTKKRHCVVRLQT 660  
 QY 661 VLERVAPTITGNLENQTTSGESIEVSCSTASGNPPPOIMWFKDNETLVEDSGIVLKDGNR 720  
 DB 661 VLERVAPTITGNLENQTTSGESIEVSCSTASGNPPPOIMWFKDNETLVEDSGIVLKDGNR 720  
 QY 721 NLTIRVRKEDGLYTCQACSVLGCACVBAFFIIIEGAQKTNLEIIILVGTAVIAMFFWL 780  
 DB 721 NLTIRVRKEDGLYTCQACSVLGCACVBAFFIIIEGAQKTNLEIIILVGTAVIAMFFWL 780  
 QY 781 LLVIIILRTVKRANGBELKTGYLSIVMDPDELPLDEHCERLPHYDASKWEPRDRKLKGLKPL 840  
 DB 781 LLVIIILRTVKRANGBELKTGYLSIVMDPDELPLDEHCERLPHYDASKWEPRDRKLKGLKPL 840  
 QY 841 GRGAFQVLEADAFGDKTATCETVAVKMLKEGATHSEHRLMSELKILIHGHILNVVN 900  
 DB 841 GRGAFQVLEADAFGDKTATCETVAVKMLKEGATHSEHRLMSELKILIHGHILNVVN 900  
 QY 901 LIGACTKPGGPLMVIIEFCFKGNLSYLSKRNFPYKTKGARFRQGDYVGAIPVDLK 960  
 DB 901 LIGACTKPGGPLMVIIEFCFKGNLSYLSKRNFPYKTKGARFRQGDYVGAIPVDLK 960  
 QY 961 RLDSITSSQSSASSGFVEEKSLSVDEEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
 DB 961 RLDSITSSQSSASSGFVEEKSLSVDEEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
 QY 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWWAPETIFDR 1080  
 DB 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWWAPETIFDR 1080  
 QY 1081 VYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRLKEGTRMRAPDYTTPEMYQTML 1140  
 DB 1081 VYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRLKEGTRMRAPDYTTPEMYQTML 1140  
 QY 1141 DCWHGSPSQRPFTSELVEHLGNLLQANAQDGKDYIVLPISETLSMEEDSGLSLTPSPVS 1200  
 DB 1141 DCWHGSPSQRPFTSELVEHLGNLLQANAQDGKDYIVLPISETLSMEEDSGLSLTPSPVS 1200  
 QY 1201 CMEEBVCCKPKHYDNTAGISQYLQNSKRSRPSVSKTFEDIPLEPEVKVIIPDDNOTDS 1260  
 DB 1201 CMEEBVCCKPKHYDNTAGISQYLQNSKRSRPSVSKTFEDIPLEPEVKVIIPDDNOTDS 1260  
 QY 1261 GMVLASEELKTLIEDRTKLSFGCMVPSKRSRSVASSEGNQTSYGYSQSHSDTDTTVYS 1320  
 DB 1261 GMVLASEELKTLIEDRTKLSFGCMVPSKRSRSVASSEGNQTSYGYSQSHSDTDTTVYS 1320  
 QY 1321 SEEAELLKLEIGVQTGSTAQILQPDGTTLSPPV 1356  
 DB 1321 SEEAELLKLEIGVQTGSTAQILQPDGTTLSPPV 1356



AC	ADH17130;	Db	181	SKKGFTIPSYMI	SVAGMVFCEAKINDES	YQIMYIVVVVGVRIYDVLS	SPSHGIELSVGE	240		
XX										
DT	11-MAR-2004 (first entry)	Qy	241	KLVLNCTARTEL	ANGIDFNWEYFSSKHQ	KLVNRDLKTQSGSEM	KKFLTLTIDGVT	300		
XX										
DE	Human vascular endothelial growth factor receptor 2 (VEGFR2) protein.	Db	241	KLVLNCTARTEL	ANGIDFNWEYFSSKHQ	KLVNRDLKTQSGSEM	KKFLTLTIDGVT	300		
XX										
KW	tyrosine kinase activity; type 1 plasminogen activator inhibitor; PAI-1;	Qy	301	DOGLYTCAASGL	MTKNSIFVRVHEK	PFVAFSGMESLVEAT	VGRRVRIPAKYLG	360		
KW	TMMP-1; tissue inhibitor of metalloproteinase 1; vinculin;									
KW	vascular endothelial growth factor; VEGF; placental growth factor; PLGF;	Db	301	DOGLYTCAASGL	MTKNSIFVRVHEK	PFVAFSGMESLVEAT	VGRRVRIPAKYLG	360		
KW	migration inhibitory factor; MIG; human;									
KW	vascular endothelial growth factor receptor 2; VEGFR2; receptor.	Qy	361	EIKWYKNGIPL	ESNHTIKAGHVLT	IMEVSRDGTGNTVILT	NPISKEKQSHVSLV	420		
XX										
OS	Homo sapiens.	Db	361	EIKWYKNGIPL	ESNHTIKAGHVLT	IMEVSRDGTGNTVILT	NPISKEKQSHVSLV	420		
XX										
PN	WO2003097854-A2.	Qy	421	POIGEKSLISP	VDSYQYGTQTTLT	CTVYAIPPPHHIHWYQ	LEBECANEP	SQAVSTNP	480	
XX										
PD	27-NOV-2003.	Db	421	POIGEKSLISP	VDSYQYGTQTTLT	CTVYAIPPPHHIHWYQ	LEBECANEP	SQAVSTNP	480	
XX										
PF	19-MAY-2003; 2003WO-US015711.	Qy	481	PCSEWSEDFQ	GNKIEVNKNQFAL	IEGKNTVSTLVIQ	AAVNSALYKCEA	VNKGGE	540	
XX										
PR	17-MAY-2002; 2003US-0380872P.	Db	481	PCSEWSEDFQ	GNKIEVNKNQFAL	IEGKNTVSTLVIQ	AAVNSALYKCEA	VNKGGE	540	
PR	24-FEB-2003; 2003US-0448874P.									
PR	24-FEB-2003; 2003US-0448922P.	Qy	541	RVISFHVTRG	PEITLQPDQMPTQ	ESVSLWCTADRST	FENLTWYKLG	PQPLPIHVGELPT	600	
XX										
XX	(SUGE-) SUGEN INC.	Db	541	RVISFHVTRG	PEITLQPDQMPTQ	ESVSLWCTADRST	FENLTWYKLG	PQPLPIHVGELPT	600	
XX										
PI	Morimoto A, Deprimo S, O'farrell A, Smolich BD, Manning WC;	Qy	601	PVCNKDLTLK	LNATMFSNSTDIL	IMELKNASLQDQD	YVCLAQDRKTK	KRHCVRQLT	660	
PI	Walter SA, Schilling JW, Cherrington J;									
PI	WPI; 2004-042604/04.	Db	601	PVCNKDLTLK	LNATMFSNSTDIL	IMELKNASLQDQD	YVCLAQDRKTK	KRHCVRQLT	660	
DR										
XX	Determining whether a test compound inhibits tyrosine kinase activity in	Qy	661	VLERVAPTIT	GNLENQTTSGES	IEVSCASGNPPQ	IMWFKDNETL	VEDSGIVLKDGNR	720	
PT	a mammal by exposing the mammal to the test compound and measuring in the									
PT	mammal the level of at least one of the measured proteins or mRNA	Db	661	VLERVAPTIT	GNLENQTTSGES	IEVSCASGNPPQ	IMWFKDNETL	VEDSGIVLKDGNR	720	
PT	transcripts.									
XX		Qy	721	NLTIRVRKE	DEGLYTCAQCS	VLGCACAKVEAFF	IEGAQEKTNLEI	ILVLGTAVIAMFWL	780	
XX										
PS	Disclosure; SEQ ID NO 129; 408pp; English.	Db	721	NLTIRVRKE	DEGLYTCAQCS	VLGCACAKVEAFF	IEGAQEKTNLEI	ILVLGTAVIAMFWL	780	
XX										
CC	The invention relates to a novel method for determining whether a test	Qy	781	LLVILIRTVK	RANGBELKTGYS	IVMDPDELPLD	DEHCERLPYD	ASKWFFPRDRKLK	840	
CC	compound inhibits tyrosine kinase activity in a mammal comprising									
CC	measuring in the mammal the level of at least one of the proteins and/or	Db	781	LLVILIRTVK	RANGBELKTGYS	IVMDPDELPLD	DEHCERLPYD	ASKWFFPRDRKLK	840	
CC	mRNA transcripts or genes for such proteins comprising type 1 plasminogen									
CC	activator inhibitor (PAI-1), TIMP-1 (tissue inhibitor of	Qy	841	GRGAFQVIT	EADAFGDKTATC	ETVAVKMLKEG	ATHSEHRLMSEL	KILTHIGHLLNVN	900	
CC	metalloproteinase 1), vinculin, vascular endothelial growth factor									
CC	(VEGF), placental growth factor (PLGF), VEGF/PLGF heterodimers or	Db	841	GRGAFQVIT	EADAFGDKTATC	ETVAVKMLKEG	ATHSEHRLMSEL	KILTHIGHLLNVN	900	
CC	migration inhibitory factor (MIG), exposing the mammal to the test									
CC	compound and then measuring in the mammal the level of at least one of	Qy	901	LLGACTKPGG	PLMVIIEFCKF	GNLSTYLRSKR	NEFVPYKTKGAR	FROCKDYVGAIPVDLK	960	
CC	the proteins and/or mRNA transcripts previously measured. The method of									
CC	the invention may be useful for determining whether a test compound	Db	901	LLGACTKPGG	PLMVIIEFCKF	GNLSTYLRSKR	NEFVPYKTKGAR	FROCKDYVGAIPVDLK	960	
CC	inhibits tyrosine kinase activity in a mammal. The current sequence is									
CC	that of the tyrosine kinase activity inhibition-related protein of the	Qy	961	RLRDSITTSQ	SSASSGFFVEK	SLSDVEEBEAP	EDLYKDFLTLEH	LICYSFQVAKGMEFLA	1020	
CC	invention.									
XX		Db	961	RLRDSITTSQ	SSASSGFFVEK	SLSDVEEBEAP	EDLYKDFLTLEH	LICYSFQVAKGMEFLA	1020	
XX	Sequence 1356 AA;									
XX	Query Match 99.9%; Score 7092; DB 8; Length 1356;	Qy	1021	SRKCIHRDLA	ARNILLSEKNVVKI	CDFGGLARDIYK	DPDYVRKGDARLPLK	WMAPETIFDR	1080	
XX	Best Local Similarity 99.9%; Pred. No. 0;									
XX	Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	Db	1021	SRKCIHRDLA	ARNILLSEKNVVKI	CDFGGLARDIYK	DPDYVRKGDARLPLK	WMAPETIFDR	1080	
Qy	1 MESKVLVALMLCVETRAASVGLPSVSLDRLPSLQIKDILTIKANTTLQITCRGQDLD 60									
Db	1 MQSKVLVALMLCVETRAASVGLPSVSLDRLPSLQIKDILTIKANTTLQITCRGQDLD 60									
Qy	61 WLFPNNQSGSEQRVEVTECSGDLFCKTLTIPIKVIQNDGTGAYKCFYRETDLASVIYVYQD 120									
Db	61 WLFPNNQSGSEQRVEVTECSGDLFCKTLTIPIKVIQNDGTGAYKCFYRETDLASVIYVYQD 120									
Qy	121 YRSPFTASVSDQHGVIYITENKNTWIPCLGSIQNLNLSLCARYPERKRPVDPGNRISWD 180									
Db	121 YRSPFTASVSDQHGVIYITENKNTWIPCLGSIQNLNLSLCARYPERKRPVDPGNRISWD 180									
Qy	181 SKKGFTIPSYMI	Qy	1261	GMVLASEEL	KTLDRKTLSP	SGMVPKSR	SRVASEG	SNQTSYG	OSGYHSD	1320
Qy	SVAGMVFCEAKINDES	Db	1261	GMVLASEEL	KTLDRKTLSP	SGMVPKSR	SRVASEG	SNQTSYG	OSGYHSD	1320

ADH17130;

11-MAR-2004 (first entry)

Human vascular endothelial growth factor receptor 2 (VEGFR2) protein.

tyrosine kinase activity; type 1 plasminogen activator inhibitor; PAI-1;

TMMP-1; tissue inhibitor of metalloproteinase 1; vinculin;

vascular endothelial growth factor; VEGF; placental growth factor; PLGF;

migration inhibitory factor; MIG; human;

vascular endothelial growth factor receptor 2; VEGFR2; receptor.

Homo sapiens.

WO2003097854-A2.

27-NOV-2003.

19-MAY-2003; 2003WO-US015711.

17-MAY-2002; 2002US-0380872P.

24-FEB-2003; 2003US-0448874P.

24-FEB-2003; 2003US-0448922P.

(SUGEN-) SUGEN INC.

Morimoto A, Deprimo S, O'farrell A, Smolich BD, Manning WC;

Walter SA, Schilling JW, Cherrington J;

WPI; 2004-042604/04.

Determining whether a test compound inhibits tyrosine kinase activity in a mammal by exposing the mammal to the test compound and measuring in the mammal the level of at least one of the measured proteins or mRNA transcripts.

Disclosure; SEQ ID NO 129; 408pp; English.

The invention relates to a novel method for determining whether a test compound inhibits tyrosine kinase activity in a mammal comprising measuring in the mammal the level of at least one of the proteins and/or mRNA transcripts or genes for such proteins comprising type 1 plasminogen activator inhibitor (PAI-1), TIMP-1 (tissue inhibitor of metalloproteinase 1), vinculin, vascular endothelial growth factor (VEGF), placental growth factor (PLGF), VEGF/PLGF heterodimers or migration inhibitory factor (MIG), exposing the mammal to the test compound and then measuring in the mammal the level of at least one of the proteins and/or mRNA transcripts previously measured. The method of the invention may be useful for determining whether a test compound inhibits tyrosine kinase activity in a mammal. The current sequence is that of the tyrosine kinase activity inhibition-related protein of the invention.

Sequence 1356 AA;

Query Match 99.9%; Score 7092; DB 8; Length 1356;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 MESKVLAVLWLCVETRAASVGLPSVSLDPLRSIQKDLITKANTTLOITCRGORDLD 60

1 MQSKVLAVLWLCVETRAASVGLPSVSLDPLRSIQKDLITKANTTLOITCRGORDLD 60

61 WLWPNNGSGEQRVEVTECSGDLFKTLTIPKVIIGNDTGAYKCFYRETDLASVIYVVD 120

61 WLWPNNGSGEQRVEVTECSGDLFKTLTIPKVIIGNDTGAYKCFYRETDLASVIYVVD 120

121 YRSPFIASVDQHGQVYITENKNTVYPCLGSTSNLNSVLCARYPKRFVPGNRIISWD 180

121 YRSPFIASVDQHGQVYITENKNTVYPCLGSTSNLNSVLCARYPKRFVPGNRIISWD 180

181 SKKGFTIPSYMI

Qy 1321 SEAEALKLIEIGVQTGSTAQILQPDGTTLSPPV 1356  
Db 1321 SEAEALKLIEIGVQTGSTAQILQPDGTTLSPPV 1356

RESULT 9  
ADQ39806  
ID ADQ39806 standard; protein; 1356 AA.

XX ADQ39806;

XX 18-NOV-2004 (first entry)

XX Human myocardial infarction-associated gene derived protein, SEQ ID 1469.

XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;  
KW cardiant; gene therapy; human.

XX Homo sapiens.

XX W02004058052-A2.

XX 15-JUL-2004.

XX 22-DEC-2003; 2003WO-US040978.

XX 20-DEC-2002; 2002US-0434778P.

XX 10-MAR-2003; 2003US-0453135P.

XX 30-APR-2003; 2003US-0466412P.

XX 23-SEP-2003; 2003US-0504955P.

XX (APPL-) APPLERA CORP.

XX Cargill M, Devlin JJ, Takoubova O;

XX WPI; 2004-533949/51.

XX N-PSDB; ADQ38978.

XX Identifying an individual who has an altered risk for developing

XX myocardial infarction by detecting a single nucleotide polymorphism in

XX the individual's nucleic acids.

XX Claim 10; SEQ ID NO 1469; 145pp; English.

XX The invention relates to a novel method for identifying an individual who  
CC has an altered risk for developing myocardial infarction. The method  
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of  
CC the nucleotide sequences given in the specification in the individual's  
CC nucleic acids, where the presence of the SNP is correlated with an  
CC altered risk for myocardial infarction in the individual. The invention  
CC further comprises: an isolated nucleic acid molecule comprising at least  
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in  
CC the specification or its complement and encoding any one of the amino  
CC acid sequences given in the specification; an isolated polypeptide  
CC comprising an amino acid sequence given in the specification; an antibody  
CC that specifically binds to the polypeptide or its antigen-binding  
CC fragment; an amplified polynucleotide containing an SNP given in the  
CC specification and which is between about 16 and 1000 nucleotides in  
CC length; a kit for detecting an SNP in a nucleic acid, comprising the  
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a  
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a  
CC method for identifying an agent useful in treating or preventing  
CC myocardial infarction. The novel detection method has cardiant activity.  
CC The nucleic acids of the invention may be used in gene therapy. The  
CC method is useful in identifying an individual who has an increased or  
CC decreased risk for developing myocardial infarction and for preparing a  
CC composition for treating or preventing myocardial infarction. This  
CC sequence represents the protein of a human myocardial infarction-  
CC associated gene containing one or more SNP's of the invention. Note: This  
CC sequence was not shown in the specification. The sequence has come from  
CC an electronic sequence listing downloaded from the WIPO website.

SQ	Sequence 1356 AA;	
Query Match	99.9%; Score 7092; DB 8; Length 1356;	
Best Local Similarity	99.9%; Pred. No. 0;	
Matches 1355; Conservative	1; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 MESKVLLAVALWLCVETRAASVGLPSVSLDPLRLSIQKQILTIKANTTQITCRGQDLD 60	
Db	1 MQSKVLLAVALWLCVETRAASVGLPSVSLDPLRLSIQKQILTIKANTTQITCRGQDLD 60	
Qy	61 WLWPNQSGSEQRVEVTECSGDLFCFKTLTIPKVIQNDTGAYKCFYRETDLASVIVYVQD 120	
Db	61 WLWPNQSGSEQRVEVTECSGDLFCFKTLTIPKVIQNDTGAYKCFYRETDLASVIVYVQD 120	
Qy	121 YRSPFTASVSDQGVVYITENKNKTVVIICLGSISNLNVSLCARYPEKRPVDPGNRISWD 180	
Db	121 YRSPFTASVSDQGVVYITENKNKTVVIICLGSISNLNVSLCARYPEKRPVDPGNRISWD 180	
Qy	181 SKKGFTIPSYMISYAGMVFCEAKINDESQSIMYIVVVVGYRIYDVVLSPSHGIELSVGE 240	
Db	181 SKKGFTIPSYMISYAGMVFCEAKINDESQSIMYIVVVVGYRIYDVVLSPSHGIELSVGE 240	
Qy	241 KLVNCTARTELNVGIDFNWEYPPSSKHQHKLVNRDLKTQSGSEMCKFSTLTIDGVTRS 300	
Db	241 KLVNCTARTELNVGIDFNWEYPPSSKHQHKLVNRDLKTQSGSEMCKFSTLTIDGVTRS 300	
Qy	301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIPAKYLGVPPP 360	
Db	301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIPAKYLGVPPP 360	
Qy	361 EIKWYKNGIPLSNHTIKAGHVLTIMEVSESDTGNVTIILTNPISEKQSHVVSLVYVVP 420	
Db	361 EIKWYKNGIPLSNHTIKAGHVLTIMEVSESDTGNVTIILTNPISEKQSHVVSLVYVVP 420	
Qy	421 PQIGEKSLIPVDVSYOYGTQTTLTCTVYAI PPHHHIHWYQWLEEECANBPSSQAVSTNYPY 480	
Db	421 PQIGEKSLIPVDVSYOYGTQTTLTCTVYAI PPHHHIHWYQWLEEECANBPSSQAVSTNYPY 480	
Qy	481 PCEEWRSVEDFOGNNKIEVKNQFALIEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE 540	
Db	481 PCEEWRSVEDFOGNNKIEVKNQFALIEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE 540	
Qy	541 RVISFHVTRGPETTLQPDMPQTEQESVSLWCTADRTSTFENLTWYKLGQPLPIHVGELEPT 600	
Db	541 RVISFHVTRGPETTLQPDMPQTEQESVSLWCTADRTSTFENLTWYKLGQPLPIHVGELEPT 600	
Qy	601 PVKCNLDTLWKNATWFSNSTNDILIMELKNASLQDQGYVCLAQDRKTKGRHCVVRLT 660	
Db	601 PVKCNLDTLWKNATWFSNSTNDILIMELKNASLQDQGYVCLAQDRKTKGRHCVVRLT 660	
Qy	661 VLERVAPTTITGNLENOTTSIGESIEVSCVCTASGNPPQIMWFKNDETFLVSDSGILVKDGNR 720	
Db	661 VLERVAPTTITGNLENOTTSIGESIEVSCVCTASGNPPQIMWFKNDETFLVSDSGILVKDGNR 720	
Qy	721 NLTIIRVRKDEGLYTCQACSVLGCACVAFPIIEGAQEKTNLEIILVGTAVIAMFFWL 780	
Db	721 NLTIIRVRKDEGLYTCQACSVLGCACVAFPIIEGAQEKTNLEIILVGTAVIAMFFWL 780	
Qy	781 LLVILIRTVKRANGBELKTYGISIVMDPELDELDEHCERLPHYDASKWEPDRRLKLGKPL 840	
Db	781 LLVILIRTVKRANGBELKTYGISIVMDPELDELDEHCERLPHYDASKWEPDRRLKLGKPL 840	
Qy	841 GRGAFQVTEADAFGIDKTATCTRTAVAKMLKGAATHSEHRLMSELKILIHGHHLNVVN 900	
Db	841 GRGAFQVTEADAFGIDKTATCTRTAVAKMLKGAATHSEHRLMSELKILIHGHHLNVVN 900	
Qy	901 LLGACTKPGGPLMWIVIEFCFKGNLSTYLRSKNEFVPYTKGARFRQGDYVGAIPVDLK 960	
Db	901 LLGACTKPGGPLMWIVIEFCFKGNLSTYLRSKNEFVPYTKGARFRQGDYVGAIPVDLK 960	
Qy	961 RRLDSTITSSQSSASSGFBVEKSLSDVEEBEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020	
Db	961 RRLDSTITSSQSSASSGFBVEKSLSDVEEBEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020	

QY 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDDYRKGDARLPLKWWAPETIFDR 1080  
DB 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDDYRKGDARLPLKWWAPETIFDR 1080  
QY 1081 VYTIQSDVWISFGVLLWEIFSLGASPYGVKIDBEFCRRLKEGTRMRAPDYTTPEMYQTM 1140  
DB 1081 VYTIQSDVWISFGVLLWEIFSLGASPYGVKIDBEFCRRLKEGTRMRAPDYTTPEMYQTM 1140  
QY 1141 DCWHGEPSPRTSELVEHNLGNLLQANAQDGKDYIVLPISSETLSMEEDSGLSLPTSPVS 1200  
DB 1141 DCWHGEPSPRTSELVEHNLGNLLQANAQDGKDYIVLPISSETLSMEEDSGLSLPTSPVS 1200  
QY 1201 CMEEVECDPKFYDNTAGISQYLONSKRKRSPVSKTFFEDIPLEBEVKVIIDDNQDTS 1260  
DB 1201 CMEEVECDPKFYDNTAGISQYLONSKRKRSPVSKTFFEDIPLEBEVKVIIDDNQDTS 1260  
QY 1261 GVLASELKTLEDRTKLSPFGMWPKSRVSGSNQTSYGSGYHSDDTDTTVYS 1320  
DB 1261 GVLASELKTLEDRTKLSPFGMWPKSRVSGSNQTSYGSGYHSDDTDTTVYS 1320  
QY 1321 SERAELKLEIGVGTGTAQILQPDGTTLSPPV 1356  
DB 1321 SERAELKLEIGVGTGTAQILQPDGTTLSPPV 1356  
RESULT 10  
ID ADQ39808 standard; protein; 1356 AA.  
AC ADQ39808;  
XX  
XX  
DT 18-NOV-2004 (first entry)  
XX Human myocardial infarction-associated gene derived protein, SEQ ID 1471.  
DE Human myocardial infarction-associated gene derived protein, SEQ ID 1471.  
XX  
XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;  
KW cardiant; gene therapy; human.  
XX  
XX Homo sapiens.  
XX W02004058052-A2.  
XX  
XX 15-JUL-2004.  
XX  
XX 22-DEC-2003; 2003WO-US040978.  
XX  
XX 20-DEC-2002; 2002US-0434778P.  
PR 10-MAR-2003; 2003US-0453135P.  
PR 30-APR-2003; 2003US-0466412P.  
PR 23-SEP-2003; 2003US-0504955P.  
XX  
XX (APPL-) APPLERA CORP.  
XX  
XX Cargill M, Devlin JJ, Iakoubova O;  
XX WPI; 2004-533949/51.  
DR N-PSDB; ADQ38980.  
DR  
XX  
XX Identifying an individual who has an altered risk for developing  
PT myocardial infarction by detecting a single nucleotide polymorphism in  
PT the individual's nucleic acids.  
XX  
XX Claim 10; SEQ ID NO 1471; 145pp; English.  
XX  
XX The invention relates to a novel method for identifying an individual who  
CC has an altered risk for developing myocardial infarction. The method  
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of  
CC the nucleotide sequences given in the specification in the individual's  
CC nucleic acids, where the presence of the SNP is correlated with an  
CC altered risk for myocardial infarction in the individual. The invention  
CC further comprises: an isolated nucleic acid molecule comprising at least  
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in

CC the specification or its complement and encoding any one of the amino  
CC acid sequences given in the specification; an isolated polypeptide  
CC comprising an amino acid sequence given in the specification; an antibody  
CC that specifically binds to the polypeptide or its antigen-binding  
CC fragment; an amplified polynucleotide containing an SNP given in the  
CC specification and which is between about 16 and 1000 nucleotides in  
CC length; a kit for detecting an SNP in a nucleic acid, comprising the  
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a  
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a  
CC method for identifying an agent useful in treating or preventing  
CC myocardial infarction. The novel detection method has cardiant activity.  
CC The nucleic acids of the invention may be used in gene therapy. The  
CC method is useful in identifying an individual who has an increased or  
CC decreased risk for developing myocardial infarction and for preparing a  
CC composition for treating or preventing myocardial infarction. This  
CC sequence represents the protein of a human myocardial infarction-  
CC associated gene containing one or more SNPs of the invention. Note: This  
CC sequence was not shown in the specification. The sequence has come from  
CC an electronic sequence listing downloaded from the WIPO website.

XX Sequence 1356 AA;

Query Match 99.9%; Score 7092; DB 8; Length 1356;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESKVLVALVLCVETRAASVGLPSVSLDPLRLSTQKQILITKANTLTQITRGQRDL 60  
DB 1 MQSKVLVALVLCVETRAASVGLPSVSLDPLRLSTQKQILITKANTLTQITRGQRDL 60  
QY 61 WLPNNQSGSEQRVEVTECSGDLFCFKTLTPKVGNDTGAYKCFYRTDLASVIVVQD 120  
DB 61 WLPNNQSGSEQRVEVTECSGDLFCFKTLTPKVGNDTGAYKCFYRTDLASVIVVQD 120  
QY 121 YRSPFIASVSDQGVVYITENKNTVIVPCLSISNLNLSLCARYPEKRFVDPGNRISWD 180  
DB 121 YRSPFIASVSDQGVVYITENKNTVIVPCLSISNLNLSLCARYPEKRFVDPGNRISWD 180  
QY 181 SKKGFTIPSYMISYAGMVFCEAKINDESYQSIYIVVVGVYRIVDVVLSHSHGIELSVGE 240  
DB 181 SKKGFTIPSYMISYAGMVFCEAKINDESYQSIYIVVVGVYRIVDVVLSHSHGIELSVGE 240  
QY 241 KLVLNCTARTELNVGIDFNWEPSSKHQHKLVNRLDKTQSGSEMKKFLTLTIDGVT 300  
DB 241 KLVLNCTARTELNVGIDFNWEPSSKHQHKLVNRLDKTQSGSEMKKFLTLTIDGVT 300  
QY 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIIPAKYLYGPPP 360  
DB 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIIPAKYLYGPPP 360  
QY 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSDRTGNVTVTILNTPISKEKQSHVSVLVVVP 420  
DB 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSDRTGNVTVTILNTPISKEKQSHVSVLVVVP 420  
QY 421 PQIGEKSLISPVDSYQYGTQTTLCTVYALPPPHHHWYQLEECANESQAVSVTNPY 480  
DB 421 PQIGEKSLISPVDSYQYGTQTTLCTVYALPPPHHHWYQLEECANESQAVSVTNPY 480  
QY 481 PCSEWSEVEDFQGNKIEVNKNQFALIEGKNKTVSTLVIQAAVNSALYKCEANVKVGRGE 540  
DB 481 PCSEWSEVEDFQGNKIEVNKNQFALIEGKNKTVSTLVIQAAVNSALYKCEANVKVGRGE 540  
QY 541 RVISFHVTRGPEITLQPDMPQTEQESVSLWCTADRSTFENLTWYKLGQPLPIHVGLPT 600  
DB 541 RVISFHVTRGPEITLQPDMPQTEQESVSLWCTADRSTFENLTWYKLGQPLPIHVGLPT 600  
QY 601 PUCKNLDLTKLNATWFSNSTNDILIMELKNASIQDQGVYVCLADRKTKKRCVVRQLT 660  
DB 601 PUCKNLDLTKLNATWFSNSTNDILIMELKNASIQDQGVYVCLADRKTKKRCVVRQLT 660  
QY 661 VLERVAPTTIGNLENOTTSTIGESIEVSCASGNPPQIMWFKDNETLVEDSGIVLKDGNR 720  
DB 661 VLERVAPTTIGNLENOTTSTIGESIEVSCASGNPPQIMWFKDNETLVEDSGIVLKDGNR 720

Qy 721 NLITRRVRKEDEGLYTCQACSVLGCACVAFVFIIEGAQEKTNLEIILVGTAVIAMFFWL 780  
Db 721 NLITRRVRKEDEGLYTCQACSVLGCACVAFVFIIEGAQEKTNLEIILVGTAVIAMFFWL 780  
Qy 781 LLVILITVRKANGELKTYLSIVMDPDELDEHCHERLPYDASKWEFPRDRKLKPKL 840  
Db 781 LLVILITVRKANGELKTYLSIVMDPDELDEHCHERLPYDASKWEFPRDRKLKPKL 840  
Qy 841 GRGAGGVIEADAFGIDKATCTAVAKMLKEGATHSEHRAALMSELKILIHGHILNVN 900  
Db 841 GRGAGGVIEADAFGIDKATCTAVAKMLKEGATHSEHRAALMSELKILIHGHILNVN 900  
Qy 901 LLGACTKPGGGLPMVIVFECFNGNLSTYLRSKRNEFPYKTKGARFRQGDYVCAIPVDLK 960  
Db 901 LLGACTKPGGGLPMVIVFECFNGNLSTYLRSKRNEFPYKTKGARFRQGDYVCAIPVDLK 960  
Qy 961 RLDSITSSOSSASSGFVEEKSLSVDEEERAPEDLYKDFLTLEHLICYSQVAKGMEFLA 1020  
Db 961 RLDSITSSOSSASSGFVEEKSLSVDEEERAPEDLYKDFLTLEHLICYSQVAKGMEFLA 1020  
Qy 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDYKPDYVRKGDARLPLKWMAPETIFDR 1080  
Db 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDYKPDYVRKGDARLPLKWMAPETIFDR 1080  
Qy 1081 VYTIQSDVMSFGVLLWEIFSLGASPGVVKIDBEFCRRLKEGTRMRAPDYTTPEMYQTML 1140  
Db 1081 VYTIQSDVMSFGVLLWEIFSLGASPGVVKIDBEFCRRLKEGTRMRAPDYTTPEMYQTML 1140  
Qy 1141 DCWHGEPSPRTFSELVEHNLGNLQANAQDGKDYIVLPISETLSMEEDSGLSLPTSPVS 1200  
Db 1141 DCWHGEPSPRTFSELVEHNLGNLQANAQDGKDYIVLPISETLSMEEDSGLSLPTSPVS 1200  
Qy 1201 CMEEEEVCDPKFHYDNTAGISQYLONSKRKRSPVSVKTFPEDIPLBEPEVKVIPPDDNQDTS 1260  
Db 1201 CMEEEEVCDPKFHYDNTAGISQYLONSKRKRSPVSVKTFPEDIPLBEPEVKVIPPDDNQDTS 1260  
Qy 1261 GMVLASELKTLEDRTKLSFGMGWPKSRRESVASEGNSQTSYGOSGVHSDDTDTTVIS 1320  
Db 1261 GMVLASELKTLEDRTKLSFGMGWPKSRRESVASEGNSQTSYGOSGVHSDDTDTTVIS 1320  
Qy 1321 SEEAELLKLEIGVQTGSTAQIILQPSGTTLSPPV 1356  
Db 1321 SEEAELLKLEIGVQTGSTAQIILQPSGTTLSPPV 1356

RESULT 11

ADR46648  
ID ADR46648 standard; protein; 1356 AA.

AC ADR46648;

DT 18-NOV-2004 (first entry)

DE Cancer-associated protein, SEQ ID 61.

KW Cytostatic; Gene Therapy; cancer; human.

OS Homo sapiens.

PN WO2004073657-A2.

PD 02-SEP-2004.

PF 19-FEB-2004; 2004WO-US005455.

PR 19-FEB-2003; 2003US-0448784P.

PA (PROT-) PROTEIN DESIGN LABS INC.

PI Aziz N, Gish KC, Wilson KE, Zlotnik A;

DR WPI; 2004-652787/63.

DR N-PSDB; ADR46590.

XX Detecting a pathological cell in a patient for diagnosing or treating  
PT cancer by detecting in a biological sample from the patient genes whose  
PS expression are up-regulated or down-regulated in specific cancers.

XX Claim 1; SEQ ID NO 61; 375bp; English.

XX The present invention relates to a method for detecting cancer in a  
CC patient. The method comprises detecting in a biological sample from the  
CC patient a nucleotide or protein sequence comprising a sequence that is at  
CC least 80% identical to a nucleotide sequence (ADR46588-ADR46645) or  
CC protein sequence (ADR46446-ADR46703). The method is useful for detecting  
CC cancer for preparing a composition for diagnosing or treating cancer.

XX Sequence 1356 AA;

Query Match 99.9%; Score 7092; DB 8; Length 1356;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESKVLALVALWLCVETRAASVGLPSVSLDLRLSLQKDLITKANTTLQITCRGORDLD 60  
Db 1 MOSKVLALVALWLCVETRAASVGLPSVSLDLRLSLQKDLITKANTTLQITCRGORDLD 60

Qy 61 WLWPNQSGSEQRVEVTECSGDLFCFKTLTPKVIGNDTGAYKCFVRETDLASVIVYVQD 120  
Db 61 WLWPNQSGSEQRVEVTECSGDLFCFKTLTPKVIGNDTGAYKCFVRETDLASVIVYVQD 120

Qy 121 YRSPFIASVSDQGVVYIITENKNTVVIPLCGISINLNSLCARYPEKRFVDPGNRISMD 180  
Db 121 YRSPFIASVSDQGVVYIITENKNTVVIPLCGISINLNSLCARYPEKRFVDPGNRISMD 180

Qy 181 SKKGFTIPSYMLSYAGWFECAKINDESQSYIMYIVVVGRIYDVVLSPSHGIELSVGE 240  
Db 181 SKKGFTIPSYMLSYAGWFECAKINDESQSYIMYIVVVGRIYDVVLSPSHGIELSVGE 240

Qy 241 KLVLNCTARTELNVGIDFNWEPSSKHQHKLVNRDLKTQSGSEMCKFLSTLITDGVTRS 300  
Db 241 KLVLNCTARTELNVGIDFNWEPSSKHQHKLVNRDLKTQSGSEMCKFLSTLITDGVTRS 300

Qy 301 DQGLYTCAASGLMTKKNSTFVRVHEKPFVAFGSGMESLVEATVGERVIRPAKILGYPPP 360  
Db 301 DQGLYTCAASGLMTKKNSTFVRVHEKPFVAFGSGMESLVEATVGERVIRPAKILGYPPP 360

Qy 361 EIKWYKNGIPLSNHTIKAGHVLTIMEVSESDTGNVTYVILTNPISEKQSHVVSVLVVYP 420  
Db 361 EIKWYKNGIPLSNHTIKAGHVLTIMEVSESDTGNVTYVILTNPISEKQSHVVSVLVVYP 420

Qy 421 POIGKSLISPVDSYQYGTQTTLCTVYAI PPPHHIHWYQLEECANEPSSQAVSVTNPY 480  
Db 421 POIGKSLISPVDSYQYGTQTTLCTVYAI PPPHHIHWYQLEECANEPSSQAVSVTNPY 480

Qy 481 PCEWRSVEDPQGNKIEVNKNQFALIEGKNTVSTLVIQAANVSALYKCEAVNKVGRGE 540  
Db 481 PCEWRSVEDPQGNKIEVNKNQFALIEGKNTVSTLVIQAANVSALYKCEAVNKVGRGE 540

Qy 541 RVISFHVTRGPETLQPDQPTQESVSLWCTADRSTFENLTWYKLGPOPLPHVGELEPT 600  
Db 541 RVISFHVTRGPETLQPDQPTQESVSLWCTADRSTFENLTWYKLGPOPLPHVGELEPT 600

Qy 601 PVCKNLDLTLWKNATMFSNSTDILIMELKNASLODQDGYVCLADQRTKKRHCVRQLT 660  
Db 601 PVCKNLDLTLWKNATMFSNSTDILIMELKNASLODQDGYVCLADQRTKKRHCVRQLT 660

Qy 661 VLERVAPITITGNLENTTSGESIEVSCCTASGNPPQIMWFKDNETFLVEDSGIVLKDGNR 720  
Db 661 VLERVAPITITGNLENTTSGESIEVSCCTASGNPPQIMWFKDNETFLVEDSGIVLKDGNR 720

Qy 721 NLTIIRVRKEDEGLYTCQACSVLGCACVAFVFIIEGAQEKTNLEIILVGTAVIAMFFWL 780  
Db 721 NLTIIRVRKEDEGLYTCQACSVLGCACVAFVFIIEGAQEKTNLEIILVGTAVIAMFFWL 780

QY 781 LLVILITVVRANGELKGTGYSIVMDPDELPLDEHCERLPYDASKWEFFPRDRILKLGKPL 840  
Db 781 LLVILITVVRANGELKGTGYSIVMDPDELPLDEHCERLPYDASKWEFFPRDRILKLGKPL 840  
QY 841 GRGAFQGVIEADAFIDKATCRTVAVKMLKEGATHSEHRAIMSELKILIHIGHILNVN 900  
Db 841 GRGAFQGVIEADAFIDKATCRTVAVKMLKEGATHSEHRAIMSELKILIHIGHILNVN 900  
QY 901 LLAGCTVGGPLMWIVFCKFNGNLSTYLRSKRNEFVPYKGFARFQGGKDYVGAIPVDLK 960  
Db 901 LLAGCTVGGPLMWIVFCKFNGNLSTYLRSKRNEFVPYKGFARFQGGKDYVGAIPVDLK 960  
QY 961 RRLDSITSSOSSASSGVBEKSLSDVEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
Db 961 RRLDSITSSOSSASSGVBEKSLSDVEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
QY 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKMAPEITFDR 1080  
Db 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKMAPEITFDR 1080  
QY 1081 VYTIQSDVMSFGVLLWEIFSLGASPPYGVKIDBEFCRRLKEGTRMRAPDYTTPEMYQTM 1140  
Db 1081 VYTIQSDVMSFGVLLWEIFSLGASPPYGVKIDBEFCRRLKEGTRMRAPDYTTPEMYQTM 1140  
QY 1141 DCWHGEPSPPTSELSVEHLGNLLOANAQDGKDIYVLPISSETLSMEEDSGLSPTSPVS 1200  
Db 1141 DCWHGEPSPPTSELSVEHLGNLLOANAQDGKDIYVLPISSETLSMEEDSGLSPTSPVS 1200  
QY 1201 CMEEVEVCDPKFYDNTAGISQYLONSKRSPVSVKTFEDIPLEBEVVKVIPPDDNQDTS 1260  
Db 1201 CMEEVEVCDPKFYDNTAGISQYLONSKRSPVSVKTFEDIPLEBEVVKVIPPDDNQDTS 1260  
QY 1261 GWVLASEELKTLIEDRTKLSFGMGVPSKRSVASESGNSQTSYGQSGYHSDDDTTTVYS 1320  
Db 1261 GWVLASEELKTLIEDRTKLSFGMGVPSKRSVASESGNSQTSYGQSGYHSDDDTTTVYS 1320  
QY 1321 SEAEALLKLEIGVGTGTAQILQPDGTTLSPPV 1356  
Db 1321 SEAEALLKLEIGVGTGTAQILQPDGTTLSPPV 1356  
RESULT 12  
ADT92353  
ID ADT92353 standard; protein; 1356 AA.  
XX AC ADT92353;  
XX DT 13-JAN-2005 (first entry)  
XX DE Human vascular endothelial growth factor receptor 2 (VEGFR2) protein.  
XX KW vascular endothelial growth factor receptor; VEGFR;  
XX KW protein co-ordinate data; X-ray diffraction; rational drug design;  
XX KW VEGFR2.  
XX OS Homo sapiens.  
XX PN WO2004092217-A1.  
XX PD 28-OCT-2004.  
XX PF 05-APR-2004; 2004WO-IB001251.  
XX PR 17-APR-2003; 2003US-0463957P.  
XX PA (PFIZ ) PFIZER INC.  
XX PI Bender SL, Kania RS, Mctigue MA, Palmer CL, Panko C;  
XX PI Wickersham J;  
XX DR WPI; 2004-784574/77.  
XX DR N-PSDB; ADT92352.

PT Crystalline structure for identifying potential vascular endothelial  
PT growth factor receptor (VEGFR) modulator, of VEGFR peptide and ligand  
XX complex.

XX Claim 4; SEQ ID NO 2; 332pp; English.

XX This invention relates to the novel crystalline structure of a vascular  
XX endothelial growth factor receptor (VEGFR) peptide-ligand complex, where  
XX the protein co-ordinate data is provided. Specifically, it refers to  
XX using X-ray diffraction techniques to provide structural information  
XX about the VEGFR kinase domain (VEGFRKD) and identify a ligand binding  
XX pocket and activation loop as defined in the specification. The present  
XX invention describes using this information to identify potential VEGFR  
XX modulators, as well as designing potential modulators using rational drug  
XX design and the three-dimensional structural information that will be able  
XX to bind to and modulate the activity of this protein. In particular, it  
XX identifies potentially important atoms and contributory amino acids by  
XX using root mean square deviation to highlight atoms that are within 1.25  
XX Angstroms of the C-alpha core. This polypeptide is the human VEGFR2  
XX protein sequence of the invention.

XX Sequence 1356 AA;

Query Match 99.9%; Score 7092; DB 8; Length 1356;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESKLLAVALWLCVETRAASVGLPSVSLDPLRLSIQKDIITIKANTLQITCRGQRLD 60

Db 1 MQSKLLAVALWLCVETRAASVGLPSVSLDPLRLSIQKDIITIKANTLQITCRGQRLD 60

QY 61 WLPNNQSGEQREVEVEVTECSGDLFCFKTLTPKVGNDTGAYKCFYRETDLASVYVTVQD 120

Db 61 WLPNNQSGEQREVEVEVTECSGDLFCFKTLTPKVGNDTGAYKCFYRETDLASVYVTVQD 120

QY 121 YRSPFTASVSDQHGVIYITENKNTVIPCGLSISNLNLSLCARYPEKRVDPGNRLSWD 180

Db 121 YRSPFTASVSDQHGVIYITENKNTVIPCGLSISNLNLSLCARYPEKRVDPGNRLSWD 180

QY 181 SKGFTIPSYMISYAGWVCEAKINDESYQSIWVIVVVGRIYDVVLSPSHGIELSVGE 240

Db 181 SKGFTIPSYMISYAGWVCEAKINDESYQSIWVIVVVGRIYDVVLSPSHGIELSVGE 240

QY 241 KLVNCTARTELNVGIDFNWEPSSKHQKLVNRLDKTQSGSEMKKFLSTLIDGVTSS 300

Db 241 KLVNCTARTELNVGIDFNWEPSSKHQKLVNRLDKTQSGSEMKKFLSTLIDGVTSS 300

QY 301 DOGLYTCASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVIRPAKYLGYPPP 360

Db 301 DOGLYTCASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVIRPAKYLGYPPP 360

QY 361 EIKWYKNGIPLSNHTIKAGHVLTIMESERDGTGNTVILTNPISKEKQSHVSVLVYVP 420

Db 361 EIKWYKNGIPLSNHTIKAGHVLTIMESERDGTGNTVILTNPISKEKQSHVSVLVYVP 420

QY 421 PQIGEKSLISPVDYQYGTQTTLCTVYAIIPPHHHIHWYQLEEECANEPSQAVSVTNPY 480

Db 421 PQIGEKSLISPVDYQYGTQTTLCTVYAIIPPHHHIHWYQLEEECANEPSQAVSVTNPY 480

QY 481 PCBEWRSVEDFQGNKIENKNOFALIEGKNKTSTLVIOAANVSALYKCEANVKYGRGE 540

Db 481 PCBEWRSVEDFQGNKIENKNOFALIEGKNKTSTLVIOAANVSALYKCEANVKYGRGE 540

QY 541 RVISFHVTVGPEITLQPDQPTBOESVSLWCTADRTFFENLTWYKLPQPLPHVGEELPT 600

Db 541 RVISFHVTVGPEITLQPDQPTBOESVSLWCTADRTFFENLTWYKLPQPLPHVGEELPT 600

QY 601 PUCKNLDTLWKLNATMFSNSTNDILIMELKNASLDQGDYVCLAQDRKTKGRHCVRQLT 660

Db 601 PUCKNLDTLWKLNATMFSNSTNDILIMELKNASLDQGDYVCLAQDRKTKGRHCVRQLT 660

QY 661 VLERVAPTTGNLENQTTSGIESIEVSCASGNPPQIMWFKONETLVEDSGVILKQGR 720

Db 661 VLERVAPTTGNLENQTTSGIESIEVSCASGNPPQIMWFKONETLVEDSGVILKQGR 720

Db 661 VLKRVAPITITGNLENQTSIGESIEVSTAGNPPQIMFMKDNELTVSDSGIIVLKDGNR 720  
 Qy 721 NLITRRVRKEDGLYTQOACSVLCAKVEAFFIIEGAQKTNLBIILVGTAVIAMFFWL 780  
 Db 721 NLITRRVRKEDGLYTQOACSVLCAKVEAFFIIEGAQKTNLBIILVGTAVIAMFFWL 780  
 Qy 781 LLVILTRVRKANGELKTYLSIVMDPDELPLDHECHERLPYDASKWEFFDRDLKLGKPL 840  
 Db 781 LLVILTRVRKANGELKTYLSIVMDPDELPLDHECHERLPYDASKWEFFDRDLKLGKPL 840  
 Qy 841 GRGAFQGVIEADAFIDKATCTRTAVKMLKEGATHSEHRAHMSLKLIIHGHILNVN 900  
 Db 841 GRGAFQGVIEADAFIDKATCTRTAVKMLKEGATHSEHRAHMSLKLIIHGHILNVN 900  
 Qy 901 LLGACTKPGGLPMLVIVFECFNGNLSTYLRKRNEFPVYKYGARFROGKDYVGAIPVDLK 960  
 Db 901 LLGACTKPGGLPMLVIVFECFNGNLSTYLRKRNEFPVYKYGARFROGKDYVGAIPVDLK 960  
 Qy 961 RLDSITSSOSSASSGFVEBKSLSDVEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
 Db 961 RLDSITSSOSSASSGFVEBKSLSDVEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
 Qy 1021 SRKCIHRDLAARNILLSEKNVWKICDFGLARDYKDPDYVRKGDARLPLKWMAPETIFDR 1080  
 Db 1021 SRKCIHRDLAARNILLSEKNVWKICDFGLARDYKDPDYVRKGDARLPLKWMAPETIFDR 1080  
 Qy 1081 VYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRLKEGTRMAPDYTTPEMYQTM 1140  
 Db 1081 VYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRLKEGTRMAPDYTTPEMYQTM 1140  
 Qy 1141 DCWHGEPQRTFSELVEHLGNLLQANAQDGDYIVLPISETLSMEEDSGLSPTSPVS 1200  
 Db 1141 DCWHGEPQRTFSELVEHLGNLLQANAQDGDYIVLPISETLSMEEDSGLSPTSPVS 1200  
 Qy 1201 CMEEVEVCDPKFYNDYNTAGISQVLQNSKRKSRPVSVKTFEDIPLBEPEVKVIPDDNQDTS 1260  
 Db 1201 CMEEVEVCDPKFYNDYNTAGISQVLQNSKRKSRPVSVKTFEDIPLBEPEVKVIPDDNQDTS 1260  
 Qy 1261 GMLVASELKLTDRTKLSFGMWPVKSRVASEGNSQTSYGQSVHSDDTDTTVYS 1320  
 Db 1261 GMLVASELKLTDRTKLSFGMWPVKSRVASEGNSQTSYGQSVHSDDTDTTVYS 1320  
 Qy 1321 SEAEALLKLIEIGVQTGSTAQILQPDGTTLSPPV 1356  
 Db 1321 SEAEALLKLIEIGVQTGSTAQILQPDGTTLSPPV 1356  
 RESULT 13  
 ADV90285  
 ID ADV90285 standard; protein; 1356 AA.  
 XX AC ADV90285;  
 XX DT 10-MAR-2005 (first entry)  
 XX DE Protease-hydrolysed polypeptide #62.  
 XX KW Protease; immune disorder; inflammation; musculoskeletal disease;  
 KW dermatological disease; gastrointestinal disease; endocrine disease;  
 KW metabolic disorder; cancer; hematological disease;  
 KW cardiovascular disease; neurological disease; neurodegenerative disease;  
 KW growth disorder; respiratory disease; genitourinary disease;  
 KW gynecological disorder; nutritional disorder; infection; cytostatic;  
 KW gastrointestinal-gen.; antiinflammatory; antidiabetic; analgesic;  
 KW antithratic; osteopathic; antidiabetic; nephrotropic;  
 KW cardiovascular-gen.; immunosuppressive; respiratory-gen.; antipsoriatic;  
 KW antiallergic; dermatological; enzyme; hydrolysis.  
 OS Homo sapiens.  
 XX W02004113522-A1.  
 XX 29-DEC-2004.

XX PF 18-JUN-2004; 2004WO-EP051173.  
 XX PR 18-JUN-2003; 2003EP-00013819.  
 PR 10-NOV-2003; 2003EP-00025851.  
 PR 11-NOV-2003; 2003EP-00025851.  
 PR 11-FEB-2004; 2004EP-00003058.  
 XX PA (DIRE-) DIREVO BIOTECH AG.  
 XX PI Haupts U, Koltermann A, Scheidig A, Voetsmeier C, Ketting U;  
 XX WPI; 2005-057985/06.  
 PT Proteases with defined specificity for a target substrate useful for  
 PT treating a specific disease related to the target substrate, such as  
 PT cancer, asthma, diabetes, inflammatory disorders and psoriasis.  
 XX Claim 50; SEQ ID NO 115; 250pp; English.  
 XX The invention relates to the use of a protease with defined specificity  
 CC for a target substrate for preparing a medicament for the treatment of a  
 CC specific disease related to the target substrate. The invention also  
 CC relates to a pharmaceutical or diagnostic composition comprising one or  
 CC more enzymes in the use cited, optionally comprising pharmaceutically or  
 CC diagnostically acceptable carriers, excipients and/or auxiliary agents, a  
 CC method for cleaving a target substrate in vivo or in vitro comprising  
 CC contacting the target substrate with a protease as cited in the use  
 CC mentioned, and a method for treatment of a disease in a patient connected  
 CC with a specific target substrate comprising administering to the patient  
 CC a protease with defined specificity for the specific target substrate.  
 CC The protease hydrolyses the target substrate and eliminates or reduces  
 CC one or more biological activities, physico-chemical properties or  
 CC pharmacological properties of the target protein and/or activates or  
 CC increases one or more biological activities, physico-chemical properties  
 CC or pharmacological properties of the target protein, and/or adds one or  
 CC more biological activities, physico-chemical properties or  
 CC pharmacological properties to the target protein. The protease may be  
 CC administered to treat immune disorders, inflammatory disorders,   
 CC musculoskeletal diseases, dermatological diseases, gastrointestinal  
 CC diseases, endocrine diseases, metabolic disorder, cancers, hematological  
 CC diseases, cardiovascular diseases, neurological diseases,  
 CC neurodegenerative diseases, growth disorders, respiratory diseases,  
 CC genitourinary diseases, gynecological disorders, nutritional disorders  
 CC and infections. This sequence represents a polypeptide hydrolysed by a  
 CC protease used in the scope of the invention.  
 XX SQ Sequence 1356 AA;  
 Query Match 99.9%; Score 7092; DB 9; Length 1356;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MESKVLVALVLMCVETRAASVGLPSVSLDPLRLSIQKDLITIKANTLIQITCRGQDLD 60  
 Db 1 MQSKVLVALVLMCVETRAASVGLPSVSLDPLRLSIQKDLITIKANTLIQITCRGQDLD 60  
 Qy 61 WLWPNQSGSEORVEVTECSGLFKTLTIPKVGNDTGAYKCFYRETDLASVIYVYVQD 120  
 Db 61 WLWPNQSGSEORVEVTECSGLFKTLTIPKVGNDTGAYKCFYRETDLASVIYVYVQD 120  
 Qy 121 YRSPFTASVSDQGVVYITENKNTVIVPCLGISNLNLSLCARYPEKFPVDPGNRISWD 180  
 Db 121 YRSPFTASVSDQGVVYITENKNTVIVPCLGISNLNLSLCARYPEKFPVDPGNRISWD 180  
 Qy 181 SKKGFTIPSYMISYAGMVFCEAKINDESYQSIMYIVVVVGYRIYDVVLSPSHGIELSVGE 240  
 Db 181 SKKGFTIPSYMISYAGMVFCEAKINDESYQSIMYIVVVVGYRIYDVVLSPSHGIELSVGE 240  
 Qy 241 KLVLNCTARTELNVGIDFNWEYPSKXKHKLNRDLKTQSGSEMCKFFLTITIDGVTWS 300  
 Db 241 KLVLNCTARTELNVGIDFNWEYPSKXKHKLNRDLKTQSGSEMCKFFLTITIDGVTWS 300

301 DOGLYTCAASGLMTKNSFTVRVHEKPFVAFSGMESLVEATVGERVRIPAKYLGYPVP 360  
301 DOGLYTCAASGLMTKNSFTVRVHEKPFVAFSGMESLVEATVGERVRIPAKYLGYPVP 360  
361 EIKWYKNGIPLESNHTIKAGHVLTIMESVSRDTGNYTVILTNPISEKQSHVSVLVVYP 420  
361 EIKWYKNGIPLESNHTIKAGHVLTIMESVSRDTGNYTVILTNPISEKQSHVSVLVVYP 420  
421 POIGESKLSIPVDSYQGTTLCTCTVYAI PPPHHIHWYQWLBEECANBESQAVSVTNPY 480  
421 POIGESKLSIPVDSYQGTTLCTCTVYAI PPPHHIHWYQWLBEECANBESQAVSVTNPY 480  
481 PCEWRVSDFOGKNLEVNKNQFALTEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE 540  
481 PCEWRVSDFOGKNLEVNKNQFALTEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE 540  
541 RVISFHVTRGPETITLOPDMQTEQESVSLMCTADRSFENLTWYKGPQLPIHVGLPT 600  
541 RVISFHVTRGPETITLOPDMQTEQESVSLMCTADRSFENLTWYKGPQLPIHVGLPT 600  
601 PVCKNLDTLKLNATMFESNSTNDILIMELKNASLODQGDYVCLADRKTKRHCVRQLT 660  
601 PVCKNLDTLKLNATMFESNSTNDILIMELKNASLODQGDYVCLADRKTKRHCVRQLT 660  
661 VLERVAPITIGNLENQTTISGESIEVSTASGNPPPOIMWFKONETLVEDSGVLKDGNR 720  
661 VLERVAPITIGNLENQTTISGESIEVSTASGNPPPOIMWFKONETLVEDSGVLKDGNR 720  
721 NLTIIRVRKEDEGLYTQACSVLGCARVEAFFIIEGAQEKTNLEIILVGTAVIANFPWL 780  
721 NLTIIRVRKEDEGLYTQACSVLGCARVEAFFIIEGAQEKTNLEIILVGTAVIANFPWL 780  
781 LLVILITVRKANGELKTYLSIVMDPDLDEHCERLPYDASKWEFPRDRKLKGLKPL 840  
781 LLVILITVRKANGELKTYLSIVMDPDLDEHCERLPYDASKWEFPRDRKLKGLKPL 840  
841 GRGAFQVIEADAFIDKTCRTVAVKMLKEGATHSEHRLMSELKILIHGHLLNVN 900  
841 GRGAFQVIEADAFIDKTCRTVAVKMLKEGATHSEHRLMSELKILIHGHLLNVN 900  
901 LLGACTKPGGLMVIWVFCFGLNLSYLRSKRNEFVPYKTKGARFROCKDYVGAIPVDLK 960  
901 LLGACTKPGGLMVIWVFCFGLNLSYLRSKRNEFVPYKTKGARFROCKDYVGAIPVDLK 960  
961 RRLDSITSSOSSASSGFVEKSLSDVEEBAPELDYKDFLTLEHLICYSFQVAKGMEFLA 1020  
961 RRLDSITSSOSSASSGFVEKSLSDVEEBAPELDYKDFLTLEHLICYSFQVAKGMEFLA 1020  
1021 SRKCIHRDLAARNILLSEKVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR 1080  
1021 SRKCIHRDLAARNILLSEKVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR 1080  
1081 VYTIQSDVWFGVLLWEI FSLGASPYGVKIDBEFCRLKEGTRMRAPDYTTPEMYOTML 1140  
1081 VYTIQSDVWFGVLLWEI FSLGASPYGVKIDBEFCRLKEGTRMRAPDYTTPEMYOTML 1140  
1141 DCWHGEPSPRTFSELVEHLNGLLQANAQDGKDYIVLPISLSEMEEDSGLSPTSPVS 1200  
1141 DCWHGEPSPRTFSELVEHLNGLLQANAQDGKDYIVLPISLSEMEEDSGLSPTSPVS 1200  
1201 CMBEVEVCDFKPHYNDTAGISQYLNKSKRSPVSVKTFEDIPLEPEVVKVIFDDNQDTS 1260  
1201 CMBEVEVCDFKPHYNDTAGISQYLNKSKRSPVSVKTFEDIPLEPEVVKVIFDDNQDTS 1260  
1261 GMVLASELKTLEDRTKLSPSFGMVPSKRESVASSEGSNOTSGYQSGYHSDDDTDTTVIS 1320  
1261 GMVLASELKTLEDRTKLSPSFGMVPSKRESVASSEGSNOTSGYQSGYHSDDDTDTTVIS 1320  
1321 SEEAELKLIEIGVQTGSTAQILOPDSGTTLSPPV 1356  
1321 SEEAELKLIEIGVQTGSTAQILOPDSGTTLSPPV 1356

## RESULT 14

ADY59385  
ID ADY59385 standard; protein; 1356 AA.

XX  
AC ADY59385;

XX  
DT 19-MAY-2005 (first entry)

XX  
DE Human VEGFR-2.

XX PDGF antagonist; VEGF antagonist; psoriasis; rheumatoid arthritis;  
KW age-related macular degeneration; antipsoriatic; dermatological disease;  
KW immune disorder; antiarthritic; antirheumatic; immune disorder;  
KW inflammation; musculoskeletal disease; age related macular degeneration;  
KW ophthalmological; ocular disease; diabetic retinopathy; antidiabetic;  
KW ophthalmological; cardiovascular disease; ocular disease.

XX Homo sapiens.

XX WO2005020972-A2.

XX 10-MAR-2005.

XX 26-AUG-2004; 2004WO-US027612.

XX 27-AUG-2003; 2003US-0498407P.

XX 28-MAR-2004; 2004US-0556837P.

XX (EYET-) EYTECH PHARM INC.

XX Shima D, Caliss P, Adamis AP;

XX WPI; 2005-214423/22.

XX N-PSDB; ADY59384.

PT Treating an ocular neovascular disorder, comprises administering a  
platelet derived growth factor antagonist and a vascular endothelial  
growth factor antagonist, simultaneously or within ninety days of each  
other.

XX Disclosure; SEQ ID NO 10; 112pp; English.

XX The invention relates to a method of treating a patient diagnosed with or  
at risk for developing a neovascular disorder, or suppressing a  
neovascular disorder in a patient. The method is useful for treating a  
patient diagnosed with or at risk for developing a neovascular disorder  
or suppressing a neovascular disorder in a patient. The neovascular  
disorder is an ocular neovascular disorder, psoriasis or rheumatoid  
arthritis. The ocular neovascular disorder is ischemic retinopathy, iris  
neovascularization, intraocular neovascularization, age-related macular  
degeneration, corneal neovascularization, retinal neovascularization,  
choroidal neovascularization, diabetic retinal ischemia, or proliferative  
diabetic retinopathy. The present sequence represents a human VEGFR  
protein.

XX Sequence 1356 AA;

Query Match 99.9%; Score 7092; DB 9; Length 1356;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESKVLVALWLCVETRAASVGLPSVSLDPLRLSTOKDILTITKANTLTQITCRGQDLD 60

DB 1 MQSKVLVALWLCVETRAASVGLPSVSLDPLRLSTOKDILTITKANTLTQITCRGQDLD 60

QY 61 WLPNNQSGSEORVEVTECSDDLFCCKTLTI PKVIGNDTGAYKCFYRETDLASVYVVD 120

DB 61 WLPNNQSGSEORVEVTECSDDLFCCKTLTI PKVIGNDTGAYKCFYRETDLASVYVVD 120

QY 121 YRSPFTASVSDQHGVIYITENKNTVVI PCLSISNLNLSLCARYPEKRFVDPGNRISWD 180

DB 121 YRSPFTASVSDQHGVIYITENKNTVVI PCLSISNLNLSLCARYPEKRFVDPGNRISWD 180







Qy 241 KLVNCTARTLNVCIDENWEPSSKHQKLVNRDLKTQSGSEKKFLSTLTDGVTRS 300  
Db 241 KLVNCTARTLNVCIDENWEPSSKHQKLVNRDLKTQSGSEKKFLSTLTDGVTRS 300  
Qy 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFGSMESLVEATVGERVRIPAKYLGYPPP 360  
Db 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFGSMESLVEATVGERVRIPAKYLGYPPP 360  
Qy 361 EIKWYKNGIPILESNHTIKAGHVLTIMEVSESDTGNVTILNPISEKQSHVSLVYVP 420  
Db 361 EIKWYKNGIPILESNHTIKAGHVLTIMEVSESDTGNVTILNPISEKQSHVSLVYVP 420  
Qy 421 POIGESLISPDVSVOYCTTQTLCTVYAIIPPHHIIHWYQLEBECANEPQAVSVTPY 480  
Db 421 POIGESLISPDVSVOYCTTQTLCTVYAIIPPHHIIHWYQLEBECANEPQAVSVTPY 480  
Qy 481 PCEWRSVEDFQGNKIEVNKNQFALIEGKNKTSTLVIQANVSALYKCEAVNKVGRGE 540  
Db 481 PCEWRSVEDFQGNKIEVNKNQFALIEGKNKTSTLVIQANVSALYKCEAVNKVGRGE 540  
Qy 541 RVISFHVTRGPEITLQPDMPTEQESVSLWCTADRSTFENLTWYKLGPPQLPIHVGLPT 600  
Db 541 RVISFHVTRGPEITLQPDMPTEQESVSLWCTADRSTFENLTWYKLGPPQLPIHVGLPT 600  
Qy 601 PCKNLDLTKLKNATMFNSNNDILIMELKNASLODQDQYVCLADQRTKRGHCVRQLT 660  
Db 601 PCKNLDLTKLKNATMFNSNNDILIMELKNASLODQDQYVCLADQRTKRGHCVRQLT 660  
Qy 661 VLERVAPITIGNLENQTTISGESIEVSTAGSNPPQIMWFKDNETLVEDSGIVLKDGNR 720  
Db 661 VLERVAPITIGNLENQTTISGESIEVSTAGSNPPQIMWFKDNETLVEDSGIVLKDGNR 720  
Qy 721 NLTIIRVRKDEGLYTQACSVLCAKVEAFFIIEGAQKTNLEIIILVGTAVIAMFFWL 780  
Db 721 NLTIIRVRKDEGLYTQACSVLCAKVEAFFIIEGAQKTNLEIIILVGTAVIAMFFWL 780  
Qy 781 LLVILITVRKANGELKTGYSIVMDDELPLDEHCEERLPYDASKWFFPRDRCLKGKPL 840  
Db 781 LLVILITVRKANGELKTGYSIVMDDELPLDEHCEERLPYDASKWFFPRDRCLKGKPL 840  
Qy 841 GRGAFGVIEADAFGIDKTATCRTVAVKMLKEGATHSEHRALMSELKILIHIGHLNVN 900  
Db 841 GRGAFGVIEADAFGIDKTATCRTVAVKMLKEGATHSEHRALMSELKILIHIGHLNVN 900  
Qy 901 LLGACTKPGGLMWIVBFCKEGLNSTYLRSKNEFVYKTKGARFRQKDYVGAIPVDLK 960  
Db 901 LLGACTKPGGLMWIVBFCKEGLNSTYLRSKNEFVYKTKGARFRQKDYVGAIPVDLK 960  
Qy 961 RRLDSITSSQSSASSGFVEEKSLSDBEEBEPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
Db 961 RRLDSITSSQSSASSGFVEEKSLSDBEEBEPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
Qy 1021 SRKCIHRDLAARNLILSEKNVVKICDGLARDIYKDPYVRKGDARLPKWMAPETIFDR 1080  
Db 1021 SRKCIHRDLAARNLILSEKNVVKICDGLARDIYKDPYVRKGDARLPKWMAPETIFDR 1080  
Qy 1081 VYTIQSDVWSFGVLLWEIPLSGSPYGVKIDEBFCRRLKEGTRMRAPDYTTPEMYQTM 1140  
Db 1081 VYTIQSDVWSFGVLLWEIPLSGSPYGVKIDEBFCRRLKEGTRMRAPDYTTPEMYQTM 1140  
Qy 1141 DCWHGEPSPRTFSELVEHLGNLQANAQQDGKDYIVLPISETLSMEEDSGLSLPTSPVS 1200  
Db 1141 DCWHGEPSPRTFSELVEHLGNLQANAQQDGKDYIVLPISETLSMEEDSGLSLPTSPVS 1200  
Qy 1201 CMEEBVCDDPKPHYDNTAGISQYLNKSRKRPVSVKTFEDIPLBEPVKVIPPDDNQDTS 1260  
Db 1201 CMEEBVCDDPKPHYDNTAGISQYLNKSRKRPVSVKTFEDIPLBEPVKVIPPDDNQDTS 1260  
Qy 1261 GMVLASELKTLEDRTKLSPSFGGWPVPSKRSVASEGNSQTSQSGYHSDDDTDTTVYS 1320  
Db 1261 GMVLASELKTLEDRTKLSPSFGGWPVPSKRSVASEGNSQTSQSGYHSDDDTDTTVYS 1320

Qy 1321 SEEAELLLKLEIGVOTGTAQILOPDSGTTLSSPPV 1356  
Db 1321 SEEAELLLKLEIGVOTGTAQILOPDSGTTLSSPPV 1356

Search completed: March 10, 2006, 18:26:46  
Job time : 153.073 secs

Page Blank (uspto)

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2006, 18:27:04 ; Search time 33.0732 Seconds  
(without alignments)  
3944.887 Million cell updates/sec

Title: US-10-633-742-6  
Perfect score: 7095  
Sequence: 1 MESKVLAVLALWLCVETRAA.....GSTAQILQDPSGTTLSSPPV 1356

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 80.\*

1: Pirl.\*

2: Pirl.\*

3: Pirl.\*

4: Pirl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7070	99.6	1356	JC1402	protein-tyrosine k
2	6100.5	86.0	1367	A41228	protein-tyrosine k
3	5031.5	70.9	1348	SJ1656	vascular endotheli
4	2981.5	42.0	1379	JC4954	protein-tyrosine k
5	2804	39.5	1363	I58375	protein-tyrosine k
6	2751.5	38.8	1298	A48999	protein-tyrosine k
7	2748.5	38.7	1338	S09982	protein-tyrosine k
8	2648	37.3	1336	I60598	Fit-1 tyrosine kin
9	2623.5	37.0	1333	I78875	receptor tyrosine
10	2570	36.2	1330	S49010	embryonic receptor
11	1218.5	17.2	1089	PFHUGA	platelet-derived g
12	1189.5	16.8	1088	1 PFRUGA	platelet-derived g
13	1184	16.7	1089	1 S33727	platelet-derived g
14	1131.5	15.9	1087	2 I51552	platelet-derived g
15	1110.5	15.7	975	2 T30816	macrophage colony-
16	1085.5	15.3	1048	2 T30816	platelet-derived g
17	1063	15.0	978	1 A49814	protein-tyrosine k
18	1057.5	14.9	960	1 JN0677	protein-tyrosine k
19	1057.5	14.9	1106	1 PFRUGB	platelet-derived g
20	1052	14.8	972	1 TVHUMD	macrophage colony-
21	1035	14.6	975	1 TVMSKT	protein-tyrosine k
22	1034.5	14.6	976	1 TVHUKT	protein-tyrosine k
23	1021	14.4	977	2 I45877	protein-tyrosine k
24	1018	14.3	1098	1 PFMGRB	platelet-derived g
25	1011.5	14.3	980	1 TVCTMD	macrophage colony-
26	999.5	14.1	954	2 I51703	c-kit-related kina
27	991.5	14.0	941	1 TVMYMD	protein-tyrosine k
28	989.5	13.9	819	1 TVCHFG	fibroblast growth
29	987.5	13.9	816	2 A49151	fibroblast growth

30	984	13.9	367	2	B56598	endothelial kinase
31	982	13.8	806	2	A35963	protein-tyrosine k
32	978	13.8	821	1	TVMSBK	fibroblast growth
33	977.5	13.8	1000	2	S18827	Flt3 protein - mou
34	976.5	13.8	814	1	A39752	fibroblast growth
35	976	13.8	822	1	TVHUGF	fibroblast growth
36	972	13.7	821	1	TVHUF2	fibroblast growth
37	972	13.7	822	2	I49289	fibroblast growth
38	970	13.7	832	2	JH0393	fibroblast growth
39	969.5	13.7	729	2	A56795	fibroblast growth
40	968	13.6	797	2	S38579	fibroblast growth
41	967	13.6	812	1	A36477	fibroblast growth
42	967	13.6	822	1	TVMSFG	fibroblast growth
43	966.5	13.6	733	2	I49293	fibroblast growth
44	966.5	13.6	823	2	B35963	protein-tyrosine k
45	963.5	13.6	822	2	A45081	fibroblast growth

#### ALIGNMENTS

##### RESULT 1

JC1402

protein-tyrosine kinase (EC 2.7.1.112) KDR - human

C;Species: Homo sapiens (man)

C;Date: 30-Sep-1993 #sequence\_revision 28-Aug-1998 #text\_change 05-Oct-2004

C;Accession: JC1402; I58357

R;Termin: B.I.; Dougher-Vermazen, M.; Carrion, M.E.; Dimitrov, D.; Armellino, D.C.; Gospo

Biochem. Biophys. Res. Commun. 187, 1579-1586, 1992

A;Title: Identification of the KDR tyrosine kinase as a receptor for vascular endothelial

A;Reference number: JC1402; MUID:93038639; PMID:1417831

A;Accession: JC1402

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-48; 'I', '50-616; 'E', '702-748; 'F', '750-1263; 'I', '1265-1290; 'P', '1292-1

A;Cross-references: UNIPROT:P35968; UNIPARC:UPI000017A3C3; EMBL:X61656; NID:G31717

R;Termin: B.I.; Carrion, M.E.; Kovacs, E.; Rasmussen, B.A.; Eddy, R.L.; Shows, T.B.

Oncogene 6, 1677-1683, 1991

A;Title: Identification of a new endothelial cell growth factor receptor tyrosine kinase

A;Reference number: I58357; MUID:92019839; PMID:1656371

A;Accession: I58357

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 3-1356 <TER2>

A;Cross-references: UNIPARC:UPI000016A991; GB:I04947; NID:G186674; PIDN:AAA59459.1; PID:G

C;Comment: This enzyme acts as a receptor for vascular endothelial cell growth factor.

C;Genetics:

A;Gene: GDB:KDR

A;Cross-references: GDB:127921; OMIM:191306

A;Map position: 4q12-4q12

C;Keywords: ATP; phosphotransferase; transmembrane protein; tyrosine-specific protein kin

F;766-790/Domain: transmembrane #status predicted <TM>

F;832-1167/Domain: protein kinase homology <KIN>

F;840-848/Region: protein kinase ATP-binding motif

Query Match 99.8%; Score 7070; DB 2; Length 1356;

Best Local Similarity 99.6%; Pred. No. 6.2e-286;

Matches 1351; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy	1	MESKVLAVLALWLCVETRAASVGLPSVLDLPRLSIQKDLITIKANTTLQITCRGQDLD	60
Db	1	MESKVLAVLALWLCVETRAASVGLPSVLDLPRLSIQKDLITIKANTTLQITCRGQDLD	60
Qy	61	WLMPNNSQSEQRVEVTECSGDLFCFKTLTIPKVIIGNDTGAYKCFYRETDLASIYVVVQD	120
Db	61	WLMPNNSQSEQRVEVTECSGDLFCFKTLTIPKVIIGNDTGAYKCFYRETDLASIYVVVQD	120
Qy	121	YRSPFIASVSDQHGVIITENKNTVVIICLGSISNLNLSLCARYPEKRPVDPGNRISWD	180
Db	121	YRSPFIASVSDQHGVIITENKNTVVIICLGSISNLNLSLCARYPEKRPVDPGNRISWD	180
Qy	181	SKKGFTIPSMYISYAGMVFCEAKINDESVOSIMYVVVGYRIYDVVLSPSHGIEISVGE	240
Db	181	SKKGFTIPSMYISYAGMVFCEAKINDESVOSIMYVVVGYRIYDVVLSPSHGIEISVGE	240

Db 181 SKKGFTIPSYMISYAGMVFCEAKINDESYOSIMYIVVVGYRIYDVVLSFSGHIELSVGE 240  
QY KLVLNCTARTELNVGIDFNWEYSSSKHOHKKLVNRDLKTQSGSEMKKFLSTLTIDGVTRS 300  
Db 241 KLVLNCTARTELNVGIDFNWEYSSSKHOHKKLVNRDLKTQSGSEMKKFLSTLTIDGVTRS 300  
QY DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIIPAKYLYGPPPP 360  
Db 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIIPAKYLYGPPPP 360  
QY EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTGNVTILNTPISKEKQSHVSVLSVYVVP 420  
Db 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTGNVTILNTPISKEKQSHVSVLSVYVVP 420  
QY PQIGKSLSPVDSYOGTQTCTCTVYATPPPHLHWTWQLEEECANEPSQAVSTNPNY 480  
Db 421 PQIGKSLSPVDSYOGTQTCTCTVYATPPPHLHWTWQLEEECANEPSQAVSTNPNY 480  
QY PCBEWRSVEDFOGKNKIEVNKNQFALIEGKNKTVSTLVIOAANVSALYKCEAVNKVGRGE 540  
Db 481 PCBEWRSVEDFOGKNKIEVNKNQFALIEGKNKTVSTLVIOAANVSALYKCEAVNKVGRGE 540  
QY RVISFHVTRGPETITLOPDMQPTQESVSLWCTADRTFFENLTWYKLGPOPLPIHVGELPT 600  
Db 541 RVISFHVTRGPETITLOPDMQPTQESVSLWCTADRTFFENLTWYKLGPOPLPIHVGELPT 600  
QY PUCKNLDITLWKNATMFSNSTDILIMELKNASLQDQGVVCLAQDRKTKRHCVVRLT 660  
Db 601 PUCKNLDITLWKNATMFSNSTDILIMELKNASLQDQGVVCLAQDRKTKRHCVVRLT 660  
QY VLERVAPITITGNLENQTTISGESIEVSCTASGNPPPPQIMWFKNDETIVSDSGIVLKDGNR 720  
Db 661 VLERVAPITITGNLENQTTISGESIEVSCTASGNPPPPQIMWFKNDETIVSDSGIVLKDGNR 720  
QY NLTIIRVRKDEGLYTCQACSVLGCACVRAFFIIEGAQKTNLEIILVGTAVIAMFFWL 780  
Db 721 NLTIIRVRKDEGLYTCQACSVLGCACVRAFFIIEGAQKTNLEIILVGTAVIAMFFWL 780  
QY LVLVILITVRKANGELKTCYLSIVMDPDELPLDEHCERLPYDASKWEPRELKLGKL 840  
Db 781 LVLVILITVRKANGELKTCYLSIVMDPDELPLDEHCERLPYDASKWEPRELKLGKL 840  
QY GRGAFQGVIEADAFGDKTATCTRTVAVKMLKEGATHSEHRALMSELKILIHIGHILNVN 900  
Db 841 GRGAFQGVIEADAFGDKTATCTRTVAVKMLKEGATHSEHRALMSELKILIHIGHILNVN 900  
QY LIGACTKPGGPLMVIIVEFCFKNLSTYLSKRNEFVPYTKGARFRQGDYVGAIPVDLK 960  
Db 901 LIGACTKPGGPLMVIIVEFCFKNLSTYLSKRNEFVPYTKGARFRQGDYVGAIPVDLK 960  
QY RLDSITSSQSSASSGFVEEKSLSVDEEAPEDLYKDFLTLEHLICYSFOVAKGMEFLA 1020  
Db 961 RLDSITSSQSSASSGFVEEKSLSVDEEAPEDLYKDFLTLEHLICYSFOVAKGMEFLA 1020  
QY SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKNWAPETIFDR 1080  
Db 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKNWAPETIFDR 1080  
QY VYTIQSDVMSFGVLLWEIFSLGASPGVKIDBEFCRRLLKEGTRMAPDYTTPEMYQTML 1140  
Db 1081 VYTIQSDVMSFGVLLWEIFSLGASPGVKIDBEFCRRLLKEGTRMAPDYTTPEMYQTML 1140  
QY DCWHGEPSPQPTFSELVEHLGNLQANAQDQGYIVLPISETLSMEEDSGLSLTPSPVS 1200  
Db 1141 DCWHGEPSPQPTFSELVEHLGNLQANAQDQGYIVLPISETLSMEEDSGLSLTPSPVS 1200  
QY CMEEEVCDPKHYDNTAGISQYLQNSKRKRSPSVSKTTFEDIPLEBPVKVIPPDDNQDTS 1260  
Db 1201 CMEEEVCDPKHYDNTAGISQYLQNSKRKRSPSVSKTTFEDIPLEBPVKVIPPDDNQDTS 1260  
QY GMVLASEELKLTEDRTKLSFGFGMVPKSRRESVASEGNSQTSYGQSHSDDTDTTVYS 1320  
Db 1261 GMVLASEELKLTEDRTKLSFGFGMVPKSRRESVASEGNSQTSYGQSHSDDTDTTVYS 1320

QY 1321 SEAEALLKLEIGVQGTGTAQILQPDSTTSSPPV 1356  
Db 1321 SEAEALLKLEIGVQGTGTAQILQPDSTTSSPPV 1356

## RESULT 2

A41228

C;Species: Mus musculus (house mouse)  
C;Date: 19-Jun-1992 #sequence revision 19-Jun-1992 #text\_change 31-Dec-2004

R;Matthews, W.; Jordan, C.T.; Gavin, M.; Jenkins, N.A.; Copeland, N.G.; Lemischka, I.R.  
proc. Natl. Acad. Sci. U.S.A. 88, 9026-9030, 1991

A;Title: A receptor tyrosine kinase cDNA isolated from a population of enriched primitive  
A;Reference number: A41228; MUID:92020984; PMID:1717955

A;Accession: A41228

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1367 &lt;MAT&gt;

A;Cross-references: UNIPROT:P35918; UNIPARC:UPI0000028D93; GB:X59397; NID:G50976; PIDN:CU  
R;Willauer, B.; Witzigmann-Voos, S.; Schmurch, H.; Martinez, R.; Moller, N.P.; Risau, W.;  
Cell 72, 835-846, 1993

A;Title: High affinity VEGF binding and developmental expression suggest Flk-1 as a major  
A;Reference number: A46065; MUID:93208880; PMID:7681362

A;Accession: A46065

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-24, 'T', '26-782, 'VL', '785-916, 'C', '918-1367 &lt;ML&gt;

A;Cross-references: UNIPARC:UPI000003CA97; GB:X70842; NID:G57923; PIDN:CAA50192.1; PID:G

A;Note: submitted to the EMBL Data Library, January 1993

R;Oelrichs, R.B.; Reid, H.H.; Bernard, O.; Ziemiacki, A.; Wilks, A.F.

Oncogene 8, 11-18, 1993

A;Title: NYK/FLK-1: a putative receptor protein tyrosine kinase isolated from E10 embryonic  
A;Reference number: I58365; MUID:93141255; PMID:8423988

A;Accession: I58365

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-678, 'D', '680-1340, 'RSPVP' &lt;OEL&gt;

A;Cross-references: UNIPARC:UPI0000170C47; GB:S53103; NID:G264004; PIDN:AAB25043.1; PID:G  
C;Genetics:

A;Gene: FLK-1; NYK

C;Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; transmembrane  
F:830-1165/Domain: protein kinase homology <KIN>

F:838-846/Region: protein kinase ATP-binding motif

Query Match 86.0%; Score 6100.5; DB 2; Length 1367;

Best Local Similarity 85.7%; Pred. No. 1e-245;

Matches 1160; Conservative 72; Mismatches 106; Indels 15; Gaps 3;

QY 1 MESKVLALVALMVCVETRAASVGLPSVLDLPRLSITQKDILITKANTTLOITCRGORDLD 60

Db 1 MESKALLAVALMVCVETRAASVGLPGLFHPKLSQKDLITLANTTLOITCRGORDLD 60

QY 61 WLWPNQSSGEORVEVTEC--SDGLFCKTLTTPKVIKNDGTGAYKCFYRDTDLASVIVVY 118

Db 61 WLWPNQRDSERVLVTECGGDSIFCKTLTTPRVVGNDDTGAYKSYRVDVASTVIVVY 120

QY 119 QDYRSPFIASVDQHGVMVITENKNTVPIPCLSISINLVSLCARYPEKRFVPCGNRIS 178

Db 121 RDYRSPFIASVDQHGIVVITENKNTVPIPCKSGISINLVSLCARYPEKRFVPCGNRIS 180

QY 179 WSKKGFTIPSYMISYAGMVFCEAKINDESYOSIMYIVVVGYRIYDVVLSFSGHIELSV 238

Db 181 WBSSETGFTLPSTYMIYAGMVFCEAKINDESYOSIMYIVVVGYRIYDVVLSFSGHIELSA 240

QY 239 GBKLVNCTARTELNVGIDFNWEYSSSKHOHKKLVNRDLKTQSGSEMKKFLSTLTIDGVT 298

Db 241 GBKLVNCTARTELNVGLDFTWHSPPSKSHHKKLVNRDVKPPPGTVAKMFLSTLTIESVT 300

QY 299 RSDQGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIIPAKYLYGYP 358

Db 299 RSDQGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIIPAKYLYGYP 358

301 KSDQGEYTCVASSGRMIKRNRTFVRVHTKPPPIAFSGMKSLVEATVGSQVRIPVKYLSYP 360  
359 PPEIKWYNGIPLESNHTIKAGHVLTIMVSEBRTGNTVILNPNPSKEQSHVSLVY 418  
361 APDIKWYNGRPIESNTMTIYVDELATIMEVTERDAGNTVILNPNPSMERQSHVSLVY 420  
419 VPOIGKSLISPDVSQYQCTTCLCTVTAIPPHIHWYMOLEECANEPQAVSVTN 478  
421 VPOIGKALISPDVSQYQCTTCLCTVTAIPPHIHWYMOLEECANEPQAVSVTN 476  
479 PYPECBWRSVEDFOGKNIEVNNQFALIEGKNKTSTLVIQAANSALYKCEAVNKVR- 538  
477 PYACKERHVEDFOGKNIEVNNQFALIEGKNKTSTLVIQAANSALYKCEAVNKVR- 536  
539 GERVISHVTRGPEITLQPMQPTQESVSLMCTADRTSTFENLTWYKLGQPIPIHVGEL 598  
537 GERVISHVTRGPEITLQPMQPTQESVSLMCTADRTSTFENLTWYKLGQPIPIHVGEL 596  
599 PTPVCKNLDLWKLNAFMSNSTNDILIMELKNASLODQGVYVCLADQDKTKRHCVVRO 658  
597 LTPVCKNLDLWKLNAFMSNSTNDILIMELKNASLODQGVYVCLADQDKTKRHCVVRO 656  
659 LTVLERVAPTITGNLENTTISIGESIEVSTAGSNPPQIMWFKDNETLVEDSGIVLKOG 718  
657 LTVLERVAPTITGNLENTTISIGESIEVSTAGSNPPQIMWFKDNETLVEDSGIVLKOG 716  
719 NRNLITRRVRKEBGLYTCQACSVLGCARVAFPIIEGAQKTNLEIILVGTAVIAMPF 778  
717 NRNLITRRVRKEBGLYTCQACSVLGCARVAFPIIEGAQKTNLEIILVGTAVIAMPF 776  
779 WLLVILRTVKRANGELKTGYSIVMDPELPLDHCERLPLVDASKWFFPRDLKLG 838  
777 WLLVILRTVKRANGELKTGYSIVMDPELPLDHCERLPLVDASKWFFPRDLKLG 836  
839 PLGRGAPQVIEADAFIGDITATCTVAVKMLKEGATHSEHRAIMSELKLIHGHILNV 898  
837 PLGRGAPQVIEADAFIGDITATCTVAVKMLKEGATHSEHRAIMSELKLIHGHILNV 896  
899 VNLGACTKPGPLWIVFECFKNLSTYLSRKNEFPVYKTKGARFQCKDYVGAIPVD 958  
897 VNLGACTKPGPLWIVFECFKNLSTYLSRKNEFPVYKTKGARFQCKDYVGAIPVD 956  
959 LKRLDSITSSQSSASSGPFVEEKSLSVDEEAPEDLYKDFLLEHLICYSFQVAKGMEF 1018  
957 LKRLDSITSSQSSASSGPFVEEKSLSVDEEAPEDLYKDFLLEHLICYSFQVAKGMEF 1016  
1019 LASRKCITHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGARLPLKWMAPETIF 1078  
1017 LASRKCITHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGARLPLKWMAPETIF 1076  
1079 DRVYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRLKEGTRMRAPDYTTPEMYQT 1138  
1077 DRVYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRLKEGTRMRAPDYTTPEMYQT 1136  
1139 MLCWCHGEPGQPTFSELVHGLNLLQANAQQQKQVIVLPISETLSMEEDSGLSPTSP 1198  
1137 MLCWCHGEPGQPTFSELVHGLNLLQANAQQQKQVIVLPISETLSMEEDSGLSPTSP 1196  
1199 VSCWEEBVCDPKPHYDNTAGISYLONSKRSRPSVKTFFEDIPLEPEPVKIPDDNOT 1258  
1197 VSCWEEBVCDPKPHYDNTAGISYLONSKRSRPSVKTFFEDIPLEPEPVKIPDDNOT 1256  
1259 DSGMVLASEELKLTLEDRTKLSPSFGMVPSKRSRESVASEGNSQTSYGQSGVHSDDTTIV 1318  
1257 DSGMVLASEELKLTLEDRTKLSPSFGMVPSKRSRESVASEGNSQTSYGQSGVHSDDTTIV 1316  
1319 YSSEAEELKLTLEIGVGTGTAQILQPDGTTL 1351  
1317 YSSEAEELKLTLEIGVGTGTAQILQPDGTTL 1340

vascular endothelial growth factor receptor 1 precursor - Japanese quail  
N;Alternate names: quail endothelial kinase 1 (Quek 1); vascular endothelial growth factor  
C;Species: Coturnix coturnix japonica (Japanese quail)  
C;Date: 07-May-1995 #sequence revision 01-Sep-1995 #text\_change 05-Oct-2004  
C;Accession: J04953; A56598; I51162; S51656  
R;Richmann, A.; Marcelle, C.; Breant, C.; Le Douarin, N.M.  
Gene 174, 3-8, 1996  
A;Title: Molecular cloning of Quek 1 and 2, two quail vascular endothelial growth factor  
A;Reference number: J04953; MUID:97017121; PMID:8863722  
A;Accession: J04953  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-1348 <E1C>  
A;Cross-references: UNIPROT:P52593; UNIPARC:UPI0000138775; EMBL:X83288; NID:G603523; PID:  
A;Note: submitted to the EMBL Data Library, December 1994  
R;Richmann, A.; Marcelle, C.; Breant, C.; Le Douarin, N.M.  
Mech. Dev. 42, 33-48, 1993  
A;Title: Two molecules related to the VEGF receptor are expressed in early endothelial ce  
A;Reference number: A56598; MUID:93378866; PMID:8396413  
A;Contents: E16 spinal cord  
A;Accession: A56598  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 910-1348 <E1C2>  
A;Cross-references: UNIPARC:UPI0000171474; GB:S65205; NID:G410680; PIDN:AAB28127.1; PID:  
A;Note: Sequence extracted from NCBI backbone (NCBIN:I37162, NCBIP:I37163)  
R;Marcelle, C.; Richmann, A.  
Oncogene 7, 2479-2487, 1992  
A;Title: Molecular cloning of a family of protein kinase genes expressed in the avian em  
A;Reference number: I50595; MUID:93056482; PMID:1281306  
A;Accession: I51162  
A;Status: preliminary; translated from GB/EMBL/DBD  
A;Molecule type: DNA  
A;Residues: 1023-1079 <MAR>  
A;Cross-references: UNIPARC:UPI00000F4F0; EMBL:X69694; NID:G395226; PIDN:CAA49364.1; PI  
A;Note: the species is not identified by the authors; the most probable species is shown  
C;Comment: This protein is an endothelial-specific receptor and binds vascular endotheli  
C;Keywords: ATP; embryo; growth factor receptor; transmembrane protein  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-1348/Product: vascular endothelial growth factor receptor 1 #status predicted <MAT>  
F;756-777/Domain: transmembrane #status predicted <TM>  
F;823-1160/Domain: protein kinase homology <KIN>  
F;831-839/Region: protein kinase ATP-binding motif

Query Match 70.9%; Score 5031.5; DB 2; Length 1348;  
Best Local Similarity 72.1%; Pred. No. 2.1e-201;  
Matches 960; Conservative 144; Mismatches 212; Indels 15; Gaps 11;

Qy 1 MESKLLVALMVCVETRAASVGLPSVSLDPLRLSIOKDLITIKANTLTITCGORDLD 60  
Db 1 MELGRLVTLVCLAPVFAGL---FISMDQPTLSIQSVLTITNDTLNITCSGQAVY 57  
Qy 61 WLWPNQSGSEORVEVTECSGLFCKTLTPKIVGNDTGAYKCFYRDTLASVYVVD 120  
Db 58 WSNPNQSSVEKRLAVTGCSEGFCKTLTLRLVIGNDTGDRCLYGLDQSOATYVVD 117  
Qy 121 YRSPFIASVDQHGVIYITENKNTVVPICGLSISNLNLSLCARYPEKRFVDPGNRISWD 180  
Db 118 YRSPFVTSVGDQGIYVIT--KNKTVVPCJGVTSNLSLHAKYPEKVPDGSISWD 175  
Qy 181 SKGFTIPSMISYAGMVFCEAKINDRESYQIMYVVVGYRIYDVVLSFSGHLSVGE 240  
Db 176 NKKGFTIPSHLINYAGMVFCEAKINDRESYQIVYVAVVGYRIYDLTMNPHYQVELAVE 235  
Qy 241 KLVLNCTARTLVNVDENWEYPSKQHKLVNRDLKTOSGSEMKKFLTLITDGVTRS 300  
Db 236 KLVLNCTVRLNVDIDFRWDYPSIKERRATI--RDLKTTAG-EIKTFVSTLIESVNLS 292  
Qy 301 DOGLYTCASSGLMTKNTFFVRVHEKPFVAFSGMESLVEATVGBRVRIPAKYLYGPPP 360  
Db 293 DKGRYTCASSGRMNKNSYFIHESPFHL-EKENVVEMKLGDTVSPVFKGYPPP 351  
Qy 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSEBRTGNTVILNPNPSKEQSHVSLVYVP 420

Db 352 EAKWYKNGKVINAHNTVKLGVALVITEATEKDAGNYTVVLTNPTNKQKRHTFTLLVNVV 411  
Qy 421 POIGSEKSLISPVDSYOGTGTOTLTCTVYVAPPPHHIHWYKOLBEECANEPQAVSVTNPY 480  
Db 412 POIGENALMPVDSYKYSQALCTTIIYVPPAAVLLWYKOLBEECTFSQKVRLGANPY 471  
Qy 481 PCBEWRSVEDFOGKNLEKVNKNOPALIEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE 540  
Db 472 ACRKWKVISEKGGNOVEI-KQVVITAGTKTVSTLVIQAAVNSALYRCMATNRAGSSE 530  
Qy 541 RVTSFHVTRGPEITLQDPMOFTQESVSLMCTADRSTFENLTVYKGGPQLPIHVGELEPT 600  
Db 531 RVISFHVTRGLEINLQPSQLTEKDNSTLQCTADKTFEKLKSWYKLSHVSQTPFGGLPM 590  
Qy 601 PUCKNLDLTWKLNATPES-NSTNDILIMELKNASLQDGDYVCLAQDRKTKKHCVCVRQL 659  
Db 591 PUCKNLDALQKLNATVSNNGENVLTLEILIRNLSLODGGDYVCIADQKKAQTHQCLVXHL 650  
Qy 660 TVLERVAPTTIGNLENTTIGESIEVSCVTSAGNPPQIMFWKDNETLVEDSGIVLKGN 719  
Db 651 TVQEPHLPRVLGNLEQNTNIGETIEVLCVNGVPPNITWFKNSETLTFEDSGIVLKGN 710  
Qy 720 RNTITRRVRKDEGLYTCQACSVLGCACVBAFFIIEGAQKTNLEIILVGTAVIAMFFW 779  
Db 711 KTLITRRVRKDEGLYTCACNLIGCKCAEAFPSVQABEKTNLEIILVGTAVIAMFFW 770  
Qy 780 LLLVILIRTVKANGGELTKGYLSIVMDPELDELDCERLKYDASKWEPFRDLKLGK 839  
Db 771 LLLVILIRTVKANGGDMTKGYLSIIMDPEVDEIDECERLKYDASKWEPFRDLKLGK 830  
Qy 840 LRGAFQGVTEADAFGIDKATCRTVAVKMLKEGATHSEHRALMSLKIILHIGHLNVV 899  
Db 831 LRGAFQGVTEADAFGIDKATCRTVAVKMLKEGATHSEHRALMSLKIILHIGHLNVV 890  
Qy 900 NLGACTKPGGPLMWIVEFCFQGNLSTYLSRKNFEPVYKTKGARFRQGD-YVGAIPVD 958  
Db 891 NLGACTKPGGPLMWIVEYCKFCGNLSAYLSRSEFIPYKMSARFRQGENYTGIDSTD 950  
Qy 959 LKERLDSITSSQSSASSGFVEEKSLSDSVEEEA-PEDLYKDFLTLEHLICYSOVAKGME 1017  
Db 951 LKQRLDSITSSQSSSTSGFVEERSLSDVEEDAGSEDLCKNPLTMBDLICYSFQVARGME 1010  
Qy 1018 FLASRKCIRDLAARNILISEKNVVKICDFGLARDYKDPDYVRKGDARLPLKWMAPETI 1077  
Db 1011 FLASRKCIRDLAARNILISDNVVKICDFGLARDYKDPDYVRKGDARLPLKWMAPETI 1070  
Qy 1078 FDRVYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRRLKEGTRMRAPDYTTPEMYQ 1137  
Db 1071 FDRVYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRRLKEGTRMRAPDYTTPEMYQ 1130  
Qy 1138 TMLDCHHGSPRSQPTSELVEHLGNLLQANAQDGKDYIVLPISETLSMEEDSGLSLPTS 1197  
Db 1131 TMLDCHHGDPKQRPTESELVEHLGNLLQANVRQDGKDYIVLPISVSLNMEEDSGLSLPTS 1190  
Qy 1198 PVSCEEEVECDPKFHYDNTAGISQYLSQKRSRPSVSKTFREDIPLERPEVKVPIDDNQ 1257  
Db 1191 PASCKEEVECDPKFHYDNTAGISQYRQGSKRSPSVSKTFREDIPL-VTVKVVQEEHQ 1249  
Qy 1258 TDSGMVLASEELKLEDRTK-LSPSPFGMVPSKRSVASEGNSQTSQYSGYSHSDDTOT 1316  
Db 1250 TDSGMVLASEELKLEEQDKQVKIPSTLAPSKSNESVMSEASNQTSQYSGYSHSDMDN 1309  
Qy 1317 TVYSSEAEALL 1327  
Db 1310 MVCSEDETELL 1320

RESULT 4  
JC4954

vascular endothelial growth factor receptor 2 precursor - Japanese quail  
N;Alternate names: Quail endothelial kinase 2; Quek 2  
C;Species: Coturnix coturnix japonica (Japanese quail)

C;Date: 31-Dec-1996 #sequence\_revision 31-Dec-1996 #text\_change 05-Oct-2004  
C;Accession: JC4954  
R;Richmann, A.; Marcelle, C.; Breant, C.; Le Douarin, N.M.  
Gene 174, 3-8, 1996  
A;Title: Molecular cloning of Quek 1 and 2, two quail vascular endothelial growth factor  
A;Reference number: JC4953; MUID:97017121; PMID:8863722  
A;Accession: JC4954  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-1379 <IC>  
A;Cross-references: UNIPARC:UPI00000FD004; EMBL:X83287; NID:g619865; PIDN:CAA58267.1; PII:  
C;Comment: This protein is an endothelial-specific receptor and binds vascular endothelial  
C;Keywords: ATP; embryo; growth factor receptor; transmembrane protein  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;789-810/Domain: transmembrane #status predicted <TMW>  
F;856-1188/Domain: protein kinase homology <KIN>  
F;864-872/Region: protein kinase ATP-binding motif

Query Match 42.0%; Score 2981.5; DB 2; Length 1379;  
Best Local Similarity 46.0%; Pred. No. 2.3e-116;  
Matches 639; Conservative 208; Mismatches 440; Indels 101; Gaps 26;

Qy 1 MSKVLVALVLMCVETRAASVGLPSVSLDLRLSLQKIDILTIKANTTIQITCRGRDLD 60  
Db 1 MKRVCTPLPLMLGIVSEADLVS--SYSMTPPLTSITEBEHIINAKDITLTITCRGQHPLE 58  
Qy 61 MWMP-----NNQ-----SGSRQVVEVTECSDDLFCCKLTIPKVINNDTG 99  
Db 59 WSWPGARWTPVGGRRRWSQPOORPVGAGNPBEDCEGTGTP--YCKVLVLTESQANDTG 116  
Qy 100 AYKCFYRETD-----LASVIYVYVQDYRSPFIASVDQGVVYIITENKNTKTVI 148  
Db 117 YHCYKYIDAKIEGTAVSAIIFVDFEQPI-----NKPEILLISKEN 162  
Qy 149 ---PCLGSTSNLNVSLCARYPEKRFVDPGNRISWSKGGFTIPSYMISVAGMVFCSEAKIN 205  
Db 163 TWVPCLVSPDLNVLTLISO--NSLIHPDRKTIFFWNKGMQVPTQLIRLSLFQCETVID 220  
Qy 206 DESYQSIMYVVVGVRIYDVVLSPSHGIELSGEKLVLNCTARTNLNVGIDPNWEPSS 265  
Db 221 NKVFSNFFIIHAGIELYDIOLYPKKAAMELVGEKVLNCTVWAEFNSGVFRPQWTPYQK 280  
Qy 266 KHQHKKLNVRLDKTQSGSEMKFSLTITDGVTRSDQGLYTCAASSGLMTKKNSTFVRVH 325  
Db 281 QMQRVIESERSLQHTELS---SILLHNVSQDGLGRYTCATNGAQLMEESTDVIH 337  
Qy 326 EKPFFVAFSGMESLVEATVG-ERVRIPAKLYGPPEIKWYKNG--IPLESNHTIKAGHV 382  
Db 338 EKPFINVEWRKGPVIEATAGDEAVKLPVKVAVPQPDFQWYKAGKLIPKQSSMQ----- 393  
Qy 383 LTIMEVSRDTCNYTVILNPISEKQSHVSVLVVVPQIGEKSLISPVDSYQVGTOT 442  
Db 394 --IKDVAEHHAGTYTLVLRNRLVGLKRLSLQLIVNVPPIHEKETSSP-SIYSRRSPQA 450  
Qy 443 LTCTVVAIPPPHHIHWYKOLBEECANEPQAVSVTNPY-----PCEWRSWSEDFQ 492  
Db 451 LTCTVYGIPAPEVIQWQRPWPCRMFRRSLNSRRAARHQDRMPECKDKWQSVRQD 510  
Qy 493 GGNKIEVNKNQFALIEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGERVIFSHVTRGP- 551  
Db 511 AVNPIESIDTWVEFVEGRNKTYSKLAIQEANYSAMKYCIAASNKVGGRDERLIFYVVTIPD 570  
Qy 552 --EITLQDPMQTEQESVSLMCTADRSTFENLTVYKGGPQLPIHVGELEPTVCKNL-DT 608  
Db 571 GFEIESQPSSEPIEGODLQLSCNADNYTYENIQWRNLNLSKLHDBEGNPLVLDCKRVHHY 630  
Qy 609 LWKLNATM-FSNSTND-ILIMELKNASLQDGDYVCLAQDRKTKKHKHCVVRQLTVLVERVA 666  
Db 631 ATKMQGELRFQDSDNATLLTIPNISLGEEDGYVCEVQNRKTRKHKCHKYISVQALEI 690  
Qy 667 PRTGNLENQTTSIGESIEVSCVTSAGNPPQIMFWKDNETLVEDSGIVLKGNRNLTI 726  
Db 691 PRLKQNLTDIWNVSDSIEMRCKVDGNHVPDISWYKDEKLVBEVSGIDLADFNQRLSIQR 750

Qy 727 VRKEDEGLYTCQACSVLGCAYVEAFPIIEGAQKTNLEIILNCTAVIAMFFWLLIIVILL 786  
 Db 751 VREEDAGLYLCSVNAKGVNCSASVSEGGDDKTINVEIIVILGTGVIIVAFWLLIIF 810  
 Qy 787 RTVRANGGELKTYGLSTVMDPDLDEHCEPLPYDASKWEFPRDRKLGKPLGRGAFG 846  
 Db 811 CNIKRPAHADIKTYGLSIIMDPGEVPLBEQCAVLPYDSSKWEFPRDRLRGLGVHGAFG 870  
 Qy 847 QVIEADAFGIDKATCTRTVAVKMLKEGATHSEHRAIMSELKILIHGHILNVNLLGACT 906  
 Db 871 KWEASAFGINKNSCETAAKMLKEGATASEQKALMSELKILIHGHILNVNLLGACT 930  
 Qy 907 KPGGLMVIIVFCKFNGHLSYLRKNEFVPYTKGARPR-QGKDYVGAIPVDLKRRLDS 965  
 Db 931 KPNGLMVIIVFCKYGNLSYLRKREGFSPYREKSPRLRQVSIYEAVRADRRSRG- 989  
 Qy 966 ITSOSASGSGFVEEKSLSVDEBEAPEDLYKDLTLLEHLICYSFQVAKGMEFLASRKC 1025  
 Db 990 --TSDSAIFNRFLMHKS-QTVQPIQEVDDLLQSPLTWEDLLICYSFQVARGMEFLASRKC 1046  
 Qy 1026 HRDLAARNILLSEKVVVKICDFGLARDIYKDPDYVRKGDARLPLKMMAPETIFDRVYTIQ 1085  
 Db 1047 HRDLAARNILLSENNVVKICDFGLARDIYKDPDYVRKGSARLPLKMMAPESIFDKVYTTQ 1106  
 Qy 1086 SDVMSFGVLLWEIPSLGASYPGVKIDBEFCRLKEGTRMRADPYTTPMYQTMDCWHG 1145  
 Db 1107 SDVMSFGVLLWEIPSLGASYPGVQINBEFCQRFKDGTRMRAPETTTAETIYRMLSCWHG 1166  
 Qy 1146 EPSORPTFSELVHLGNLQANQDGYIVLPISITLMEEDSGLSLPTSPVSCMEBE 1205  
 Db 1167 DPKEPTFSDLVLEILGNLQENVOQEGKDYI--PLNDSHSEDDGFSQVSSAQONSDEE 1224  
 Qy 1206 EVCDPKF-----HYDNTAGISOLQNSKRKSPVSVKTFEDIPLEPEVKYIPDNDQT 1258  
 Db 1225 DP-DMRIRCHSLAARYNCVSFPCLTGGQINRCSSRIKTFEFPMTHTMYKAHP-DNQT 1282  
 Qy 1259 DSGMVLASELKTLEDRTKLSPSFGMVPSKRSRESVASEGNSQTSYGVSQGHSDDTTIV 1318  
 Db 1283 DSGMVLASEBFERTENRHKRGFGFSKGPNTAE-LSAEGSDLRGRCPRSYSGVQGGTF 1341  
 Qy 1319 YSSEAEEL 1326  
 Db 1342 YNSEYEL 1349

RESULT 5  
 158375  
 protein-tyrosine kinase (EC 2.7.1.112) flt4 - mouse  
 C;Species: Mus musculus (house mouse)  
 C;Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 05-Oct-2004  
 C;Accession: 158375; B42010  
 R;Finnerty, H.; Kelleher, K.; Morris, G.E.; Bean, K.; Merberg, D.M.; Kriz, R.; Morris, J.  
 Oncogene 8, 2293-2298, 1993  
 A;Title: Molecular cloning of murine FLT and FLT4.  
 A;Reference number: 158375; MUID:9330572; PMID:8393164  
 A;Accession: 158375  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-1363 <RES>  
 A;Cross-references: UNIPROT:P35917; UNIPARC:UPI0000028D92; GB:L07296; NID:9293780; PIDN:  
 R;Galland, F.; Karamysheva, A.; Mattei, M.G.; Rosnet, O.; Marchetto, S.; Birnbaum, D.  
 Genomics 13, 475-478, 1992  
 A;Title: Chromosomal localization of FLT4, a novel receptor-tyrosine kinase gene.  
 A;Reference number: A42010; MUID:92307693; PMID:1319394  
 A;Accession: B42010  
 A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra  
 A;Molecule type: DNA  
 A;Residues: 1033-1072 <GAL>  
 A;Cross-references: UNIPARC:UPI000017A423  
 C;Genetics:  
 A;Gene: FLT4  
 C;Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase

F;843-1176/Domain: protein kinase homology <KIN>  
 F;851-859/Region: protein kinase ATP-binding motif

Query Match 39.5%; Score 2804; DB 2; Length 1363;  
 Best Local Similarity 44.1%; Pred. No. 5.2e-109;  
 Matches 610; Conservative 207; Mismatches 465; Indels 102; Gaps 29;

Qy 1 MESKVLAVLALMVCVETRAASVGLPS-VSLDLRLSLQDKILITKANTTLOITCRGQDL 59  
 Db 1 MQPGAALNRLMLCL---GLLOGLANGSYMTPTLNITEDSYVIDTGDLSLSISCRGHPL 57  
 Qy 60 DWLWPNNQ-----SGSEQRVEVTECSDEL---PCKTLTIKPIKVIQNDGTGAYKCFYR-- 106  
 Db 58 EWTWPGAQEVLTTCGKSDSEDRVHDC-EGTEARPYCKVLLLAQTHANNTGSHCYKYI 116  
 Qy 107 ----ETDLASVIYVQDYRSPFIASVDQGVVYITENKNTKTVVIVPCLSGISNLNLSLC 162  
 Db 117 KARIEGTTAASTVYVVRDPKHPINKPDT-----LLVNRKDSMWVPCVSLPGLNITL- 169  
 Qy 163 ARYPEKRFVDPGNRIKSWDSKGTIISYMSIYAGMVFCEAKINDESYQSIYIVVVGYR 222  
 Db 170 -RSQSALHPDGOEVLWDDRRGRVPTQLLRDALYLQCTETWGDQNFSLNVVHITGNE 228  
 Qy 223 IYDVLSPSHGIELSVEKLVNCTARTLNVGIDFNWEYPSKSHQHKLVNRDLKTQSG 282  
 Db 229 LYDQLYPKKSMELLGKLVNCTVWAEPSDGVTFDWDYPGQAERAKVPPERSQOHT 288  
 Qy 283 SEMKFLSTLTIDGVTRSDQGLYCAASSGLMTKQNSTFVRVHEKPFVAFSGMESLVEA 342  
 Db 289 TELS---SILTIHNVQNDLGPVVCANNGIQRFRESTEVIIVHEKPFISVEMWLGKPVLEA 345  
 Qy 343 TVG-ERVRIPAKYLGTPPPPIKWKYKGIPLSHHTIKAGHVLIMEVSEEDTNGYTVILT 401  
 Db 346 TAGDELVKLPKLAAYPPPEFYQYDKRKAIVTGRHN---PHALVLKEVTEASAGVYTLALW 402  
 Qy 402 NPISKEKSHVSLVVPVPOICEKSLISPVDSYQVGTCTCTCTVVAIPPHHHHWYQ 461  
 Db 403 NSAGLRQNLISLELVNVPPIHKEASSP-SIYSHRSRQTLTCTAYGVQPLSVQHWHR 461  
 Qy 462 LEEEC---ANEPSQAVSVTNPY-CEEWSVEDFQGGNKIEVNKNQFALIEGKNKTVSTL 517  
 Db 462 PWTPTCTFAQRSLRRRQQRDGMPCQCRDKWEVTTQDAVNPLESLSWTEFVEGKNKTVSKL 521  
 Qy 518 VIOAANVSALYKCEAVNKVGRGERSVIFHTVQPE---ITLQDPMQTEGESVSLMCTAD 574  
 Db 522 VIOQANVSAMKCVVNVKVGQDERLIYFYVTTIPDGFSIESEPSDEPLEGQVRLSCRAD 581  
 Qy 575 RSTFENLTWYKLGQPLPIHVGEPLTPVCKNLTLMKLNATMFSNSTD-----ILI 626  
 Db 582 NYTYEHLRWYRLNLTSLHDAQGNPLLLDCKNV-----HLFATPLEANLEEAEPGARHATLS 637  
 Qy 627 MELKNASLOQGDYVYCLAQDRKTKGRHCVVVRQLTVLERVAPTITGNLENQTTSIGESIEV 686  
 Db 638 LNIIPVAPDEGDYVCEVQDRSQDKHKYKLSVQALEAPRLTQNLTLNVLNVSLSLEM 697  
 Qy 687 SCTASGNPPQIMFWKDNETHLVEDSGIVLQKGNRNLTIIRVKEKDEGLYTCQACSVLGCA 746  
 Db 698 RCPVAGAHVPSIVVYKDERLLEKESGIDLADSNQRLSIQRVREBEDAGRYLCSVCNAGCV 757  
 Qy 747 KVAFPIIEGAQKTNLEIILNCTAVIAMFFWLLIIVILLRTVKRANGELKTYGLSIVM 806  
 Db 758 NSSASVAVESGSDKSGMEIIVILGTGVIIVAFWLLIIFCNMKRPAHADIKTYGLSILM 817  
 Qy 807 DPDELPLDEHCERLPYDASKWEFPRDRKLGKPLGRGAFQGVIEADAFGIDKATCTRTVA 866  
 Db 818 DPGEVPLEEQCYLSYDASQWEPFRRLHGRVLGHGAFKGVVEASAFGINKSGSCDTVA 877  
 Qy 867 VKMLKEGATHSEHRAIMSELKILIHGHILNVNLLGACTKPGGLMVIIVFCKFNGHLS 926  
 Db 878 VKMLKEGATHSEHRAIMSELKILIHGHILNVNLLGACTKPNGLMVIIVFCKYGNLSN 937  
 Qy 927 YLRKNEFVPYTKG-----ARFQKQDYVGAIPVDLKRRLDSITSSQSASSGFFVEKS 982

Db 938 FLVRKTRDTFNPAEKSPEQRRRFR---MVEGAKADRRP-----GSSDRALFRFLMGKG 990  
QY 983 LSDVEEERAP-----EDLYKDFLTLEHLICYSFOVAKGMEFLASRKICHRDLAARNILLS 1037  
Db 991 ----SARRAPLVQEAEDLWLSPLTWEDLVYCFVARGMEFLASRKICHRDLAARNILLS 1046  
QY 1038 EKNVVKICDFGLARDIYKDPYVRKGDARLPLKWMAPETIDFDRVYTTQSDVWFGVLLWE 1097  
Db 1047 ESDIVKICDFGLARDIYKDPYVRKGSARLPLKWMAPESIDFKVYTTQSDVWFGVLLWE 1106  
QY 1098 IFSLGASPYGVKIDFECRLKGTMRAPDYTTPEMYQTMCLDCHWGFSPQRTSELY 1157  
Db 1107 IFSLGASPYGVQINEEFQCRKLDGTRRAPLAPAIIRHIMQSCWSDGPKARPAFSDLV 1166  
QY 1158 EHLGNLQANA-QODGKDYIVLPISETLSMEEDSGLSLTPSPVSCME-BEEVCDPKFH-- 1213  
Db 1167 EILGDLLOGGWQEERVALHSSQ--SSEEDGFMQASTALHITADADDSPPSMCH 1224  
QY 1214 -----YDNTAGISQYLQNSKRKSRPVSVKTFEDIPLBEPVKVIPPDDNQDSDGMVLASEE 1268  
Db 1225 SLAARYNCVSPFGRLARGTKTPGSSRMKTFEELP-TPITYKASMDNQDSDGMVLASEE 1283  
QY 1269 LKILEDRTKLSPFGGMVPSKSRVASEGNSQ-----TSGVQSGVHSDDTTTVYSSE 1322  
Db 1284 FEELSHRPEGSFCKPGQGHMDIPRGHPDPQGRRRRTQGAQGG-----KVFPYNE 1336  
QY 1323 EDEL 1326  
Db 1337 YGEV 1340

RESULT 6  
A48999  
N;Alternate names: class III receptor tyrosine kinase FLT4; fms-like tyrosine kinase 4  
C;Species: Homo sapiens (man)  
C;Date: 21-Jan-1994 #sequence, revision 18-Nov-1994 #text\_change 31-Dec-2004  
R;Accession: A48999; G02316; S36130; A42010  
R;Apurlikova, K.; Apurlikova, O.; Korhonen, J.; Kaipainen, A.; Pertovaara, L.; Alitalo, R.;  
Cancer Res. 52, 5738-5743, 1992  
A;Title: FLT4 receptor tyrosine kinase contains seven immunoglobulin-like loops and is e  
A;Reference number: A48999; MUID:93007958; PMID:1327515  
A;Accession: A48999  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-1298 <PAJ>  
A;Cross-references: UNIPROT:P35916; UNIPARC:UPI0000138777; PIDN:AAB23636.1; PID:G257352  
A;Experimental source: HEL erythroleukemia cells  
A;Note: sequence extracted from NCBI backbone (NCBIP:115335)  
R;Apurlikova, O.; Pajusola, K.; Partanen, J.; Armstrong, E.; Alitalo, R.; Bailey, S.K.;  
Cancer Res. 52, 746-748, 1992  
A;Title: FLT4, a novel class III receptor tyrosine kinase in chromosome 5q33-qter.  
A;Reference number: A4930; MUID:92119639; PMID:1310071  
A;Accession: A4930  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 761-1190 <APR>  
A;Cross-references: UNIPARC:UPI000017A3EA; GB:X68203; NID:G31433  
A;Note: sequence extracted from NCBI backbone (NCBIP:78155)  
R;Wood, W.I.  
submitted to the EMBL Data Library, December 1995  
A;Reference number: H01039  
A;Accession: G02316  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1298 <WO>  
A;Cross-references: UNIPARC:UPI0000138777; EMBL:U43143; NID:g1150990; PIDN:AAA85215.1; H  
R;Galland, F.  
submitted to the EMBL Data Library, December 1992  
A;Reference number: S36130  
A;Accession: S36130  
A;Molecule type: mRNA  
A;Residues: 1-23, 'D', '25-744, 'P', '746-751, 'RP', '754-889, 'Q', '891-1127, 'V', '1129-1145, 'H', '1147

A;Cross-references: UNIPARC:UPI000003CA99; EMBL:X69878; NID:g297049; PIDN:CAA49505.1; P11  
R;Galland, F.; Karamysheva, A.; Mattei, M.G.; Rosnet, O.; Marchetto, S.; Birnbaum, D.  
Genomics 13, 475-478, 1992  
A;Title: Chromosomal localization of FLT4, a novel receptor-type tyrosine kinase gene.  
A;Reference number: A42010; MUID:92307693; PMID:1319394  
A;Accession: A42010  
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual trar  
A;Molecule type: mRNA  
A;Residues: 776-889, 'Q', '891-1127, 'V', '1129-1145, 'H', '1147-1163, 'D', '1165-1200 <GAL2>  
A;Cross-references: UNIPARC:UPI000017A3EB  
C;Genetics:  
A;Gene: GDB:FLT4  
A;Cross-references: GDB:128732; OMIM:136352  
A;Map position: 5q34-5q35  
C;Keywords: ATP: autophosphorylation; phosphoprotein; phosphotransferase; transmembrane  
F;1-23/Domain: signal sequence #status predicted <SIG>  
F;24-1298/Product: protein-tyrosine kinase FLT4 #status predicted <MAT>  
F;843-1176/Domain: protein kinase homology <KIN>  
F;851-859/Region: protein kinase ATP-binding motif

Query Match 38.8%; Score 2751.5; DB 2; Length 1298;  
Best Local Similarity 44.9%; Pred. No. 7.4e-107;  
Matches 597; Conservative 195; Mismatches 458; Indels 79; Gaps 24;

QY 1 MESKVLVALMLCVETRAASVGLPSVSLDPLRLSIQKDIITIKANTTLOITCRGORDLD 60  
Db 1 MQRGAALCLRLMLCLGLLDGLVS--CYSWTPPTLNITEESHVDTGDSLSISCRGQHPLE 58  
QY 61 WLWPNQSG-----SEORVEVTECSGL-----FKTTLTPKVIQNDTGAYKCFR--- 106  
Db 59 WAWPGAQEA PATGDKDSED TGVRVDC-EGTDARPYCKVLLHVEHANDTGSVYCYKIK 117  
QY 107 ---ETDLASVIVYVQDYSRPTASVSDQHGVVVITENKNTVVPCLGISNLNLSLCA 163  
Db 118 ARIEGTAASSVYVDFDFQPFINKPDT-----LVLNRDANWVCLVSPGLNVL-- 169  
QY 164 RYPERKFEVDPGNNRISWDSKKGFTIPSMYSIAGWVCEAKINDIESQSIYIVVVVGYRI 223  
Db 170 RSQSSVLWPDGQEVVWDRRGLMVSPLLHDALYLQCEITWGDQDFLSNPFLVHITGNEL 229  
QY 224 YDVVLSPSHGIELSVGEKVLNCTARTELNVLGIDFWPEYSSKHQHKLVNRLDLTKQS 283  
Db 230 YDQLLPKRSLELLVGEKVLNCTWAEFNSGVTFDWDYFGQAERGVKVPERRSQTH 289  
QY 284 EMKKFLSTLTIDGVTESDQGLYTCASSGLMTKNSTFVRVHEKPFVAFSGSMESLVEAT 343  
Db 290 ELS---SILTIHNVSDGLSGVCKRANNGIQRFRETEFVIVHENPFISVENLKGPILEAT 346  
QY 344 VG-ERVIRIPAKYLGYPPEIKWYKNGIPILESNHTIKAGHVLTIMEYSEBDTNGYTVILTN 402  
Db 347 AGDELVKLPVLAAYPPPEFQWYKDGKALSGRH---SPHALVLKVEATEASTGYTLALWN 403  
QY 403 PISKEQSHVSLVYVVPQIGEKSLISPVDSYQYGTQTTLCTVYVAIPPHIHHYWQL 462  
Db 404 SAAGLRNRISLELVNVPQIHEKEASSP-SIYSRHSRQALTCTAYGVPLPLSIQHWHP 462  
QY 463 EEEC---ANEPSQAVSVTPNP-CEEWSVEDFQGGNKIEVNKNQFALLTEGKNKTSTVLV 518  
Db 463 WTPCKMFAORSRRRQQDMLQCRDRAVTTQDAVNPIESLDTWTFEFGKNKTYSKLV 522  
QY 519 IQAANVSALYKCAEVNKGGRVTSFHVTRGPE---ITLQPDMPQTEQSSVSLWCTADR 575  
Db 523 IQNANVSAMYKCVSNKVGDERLIYFVYTTIPDGTIESKPESELELQGPVLLSQADS 582  
QY 576 STFENLTWYKLGPOPLPIHVGELPTPVCKNLDLWKLNATMFSNSNDI-----LIM 627  
Db 583 YKYEHLRWYRLSLTLHDAGNPLLLDCKNV---HLFATPLAASLEEVAPGARHATLSL 638  
QY 628 ELKNASLQDQGVCLAQDRKTKKRCVVRQLTVLERVAPTITGNLENQTTSGEISVS 687  
Db 639 SIPRVAPEHEGHVCEVQDRRSHDKCHKYLSVQMALEAPRTQNTDLLVNDSLEMQ 698  
QY 688 CTASGNPPPPQIMWFKDNETLVSDSGILVKDGNRNLTIIRVRKDEGLYTCQACSVLGC 747



Db 699 CLVAGAHAPSIVWYKDERLLEESGVDLADSNQKLSIORVEEDAGRYLCSVNAKGCVN 758  
Qy 748 VEAFPIIEGAQKTNLEIILVGTAVIAMPFWLLVILTRVIRANGGELKTGYLSIVMD 807  
Db 759 SSASVAVEGSEDKSMIEVILVGTAVIAMPFWLLVILTRVIRANGGELKTGYLSIVMD 818  
Qy 808 PDELPLDEHCERLPYDASKWEFFDRDLKIKPILGRGAFQGVIEADAGIDKATCTRTAV 867  
Db 819 PGEVLPEQCEYLSYDASQWEPFRERHLGRVLGVGAFKQVVEASAFGIHKGSSCTVAV 878  
Qy 868 KMLKEGATHSEHRALMSELKILIHGHHLNVNLLGACTKPGGLMIVIVFCKFGNLSY 927  
Db 879 KMLKEGATHSEHRALMSELKILIHGHHLNVNLLGACTKPGGLMIVIVFCKFGNLSN 938  
Qy 928 LRSKNEFVPIYKVG----ARFQGGDYVGAIPVDLKR--RLDSITSQSSASSGFFVEEK 981  
Db 939 LRAKRDAPSPCAEKSPQRGRFRA--MVLEALDRRRPSSDRVLPAFSPKTEGGARRA 995  
Qy 982 SLSDVEEAPEDLYKDFLTHLEHLYCYSFQVAKGMEFLASRKCIRHDLAARNILLSEKNV 1041  
Db 996 S-----PDEA-EDLWLSPLTMEDELVCYSFQVARGMEFLASRKCIRHDLAARNILLSESDV 1050  
Qy 1042 VKICDFGLARDIYKDPYVRKGDARLPLKWNAPETIEDRVVTIQSDVWSFGVLLWEIFSL 1101  
Db 1051 VKICDFGLARDIYKDPYVRKGSARLPLKWNAPESIFDKVYTTQSDVWSFGVLLWEIFSL 1110  
Qy 1102 GASPYGVKIDEEFCRLKEGTRMAPDYTPPEMYOTMLDCWHGEPSPQRTFSELVEHLG 1161  
Db 1111 GASPYGVQINEPFCRLRGTMRAPELATPAIRRLMNCWSGDKPAPFASSELVEILG 1170  
Qy 1162 NLQANNAQDGKQYIVLPISETLSMEEDSGLSLPTSPVSCMBEEVCDP-----KHF 1213  
Db 1171 DLLQGRGLOBEEVCMAPRS--SOSSESGFSQVSTWALHIAQADAEDSPSLQRHSLAAR 1229  
Qy 1214 YDNTAGISQVQLNSKRSPVSVKTFREDIPLERPEVKVIPPDDNQTSQGMVLASELKTLE 1273  
Db 1230 YNWVSPGCLARGAETRGSRMKTPEEFPM--PTTYKGSVDNQDQSGMVLASEEPEQIE 1288  
Qy 1274 DRTKLSPSF 1282  
Db 1289 SRHQESGF 1297

RESULT 7  
S09982  
N;protein-tyrosine kinase (EC 2.7.1.112) flt1 precursor - human  
C;Species: Homo sapiens (man)  
C;Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 05-Oct-2004  
C;Accession: S09982  
R;Shibuya, M.; Yamaguchi, S.; Yamane, A.; Ikeda, T.; Tojo, A.; Matsushima, H.; Sato, M.  
Oncogene 5, 519-524, 1990  
A;Title: Nucleotide sequence and expression of a novel human receptor-type tyrosine kinase  
A;Reference number: S09982; MUID:90221591; PMID:2158038  
A;Molecule type: mRNA  
A;Residues: 1-1338 <SHI>  
A;Cross-references: UNIPROT:P17948; UNIPARC:UPI0000138773; EMBL:X51602; NID:G31431; PIDN  
C;Genetics:  
A;Gene: GDB:FLIT1  
A;Cross-references: GDB:120616; OMIM:165070  
A;Map position: 13q12-13q12  
C;Keywords: ATP; autophosphorylation; glycoprotein; membrane protein; phosphoprotein; ph  
F;1-22/Domain: signal sequence #status predicted <SIG>  
F;23-1338/Product: protein-tyrosine kinase flt #status predicted <MAT>  
F;825-1161/Domain: protein kinase homology <KIN>  
F;833-841/Region: protein kinase ATP-binding motif

Query Match 38.78; Score 2748.5; DB 2; Length 1338;  
Best Local Similarity 44.38; Pred. No. 1e-106;  
Matches 608; Conservative 212; Mismatches 436; Indels 115; Gaps 30;

Qy 6 LLAVALWLCVETRAASVGLPSVSLDLPRLSIQKIDILTIKANTTLOITCRGORDLWLWPN 65  
Db 9 VLICALLSCLLLTGSSSG----SKLKDPELSLKTQHIHQAGQTLHQCRGAHAKHSLPE 65  
Qy 66 NQSGSQRVEVTECSBG----LFCKTLTIPKVINQDGTAYKCFY-----RETDLASVIY 115  
Db 66 MVSKESERLSITKSACGRNGKQFCSTLTNTAQANHTGFYSCKYLAVPTSKKXETESAIY 125  
Qy 116 VYQDVRSPPIASVSQHQGVVYITENKNTVITPCIGSISNLNVSLCARYPEKRFVDPGN 175  
Db 126 IFISDTCRPEVMEYSIPEIHHTEGRE--LVTPCRVTSNITVTL--KKFPLDTLIDPGK 182  
Qy 176 RISWDSKGGFTIIPSYMISVAGMVFCAKINDESVQSIMYIVVGVYIYDVWLSPSHGIE 235  
Db 183 RIITWDSRKGFIISNATYKEIGLITCAATVNGHYLKT--NYLTHRQTNTIIDVQISTPRPVK 241  
Qy 236 LSVGEKLVNCTARTALNVGIDFNWEPYSSKHQKLVNRDLTKTQSGEMKKFLSTLTID 295  
Db 242 LLRGHTLVNCTATTPLNTRVQMTWSYPDEKKNRASVRRR--IDQSNSHANIFYSVLTID 299  
Qy 296 GVTRSDOGLYTCAASSGLMTKONSTFVRVHEKPFVAFGSGMESLVEATVGER--VRIPAKY 354  
Db 300 KMQNKDKGLYTCRVRSGSPFSKSVNTSVHIYDKAFITVXHKRQQLVETVAGRSYRLSMKV 359  
Qy 355 LGYPPPEIKWYKNGIPL--ESNHTIKAGHVLTITMEYSEKDTGNYVTLLTNPISKEKQSHV 412  
Db 360 KAPFSEVVMKDGLPATEKARYLTRGYSLLIKDVTEDAGNYTILLSIKQSNVFNKLT 419  
Qy 413 VSLVWVVPQIIGSKSLISPVDS--YQVGTQTITCTVYAIPPPPIHHWYQWLBECANEP 470  
Db 420 ATLIVNVKQIYEKAVSSPDPALYPLGSKQILCTAYGIQOP--TIKFW--HPCNNHH 475  
Qy 471 SQAVSVNTPYPCBEWSVE-----DFQGNKIEVNKNQFALIEGKNKTVSTLIVIOAAN 523  
Db 476 SEA-----RCDFCSNNEESFILDADSNMGNRIESTORMAIEGKNKMASTILVADSR 528  
Qy 524 VSALYKCEAVNKVGRERVISFHVTRGPE--ITLQPDMPQTEQESVSLWCTADSTFENLT 582  
Db 529 ISGIYICIASNKVGTVGRNISFYITDVPNGFHVNLKMPTEGEBDLKLSCTVNFPLYRDTV 588  
Qy 583 W-----YKLGPOPLPIHVGEPLTPVCKNLDLTMLKNATMFSNNDILIMELK 630  
Db 589 WILLRTVNNRTMHSYSKQ-----KMAITKEHSITLNTIM-- 624  
Qy 631 NASLQDQGDYVCLAQDRKTKRHCVVVRQLTVLSEVAPITIGNLENOTTSIGSEIVSCTA 690  
Db 625 NVSLQDSGTVACRARNVYTGEEILQKEITIRQDEAPYLRLNLSHDHTVAISSSTLIDCHA 684  
Qy 691 SGNPPQIMWFKDNETILVEDSGIVLKDGNRLTIRVRKDEGLYTCQACSVLGCACVEA 750  
Db 685 NGVPEPQITWFKNNHIQOEPGIIILGPGSSTLFIERVTEDEGVYHCKATNQKGSVESSA 744  
Qy 751 FFIIEGAQKTNLEIILVGTAVIAMPFWLLVILTRVIRANGGELKTGYLSIVMDPDE 810  
Db 745 YLTQGTSDKSNLELITCTVAAITLFWLLLLTLIRKMKRSS--SEIKTDYLSIIMDPDE 803  
Qy 811 LPLDEHCERLPYDASKWEFFDRDLKIKPILGRGAFQGVIEADAGIDKATCTRTAVAKML 870  
Db 804 VPLDEQERLPHYDASKWEFARERLKLKSLGRGAFGVQVQASAFGIKSPCTRTAVAKML 863  
Qy 871 KEGATHSEHRALMSELKILIHGHHLNVNLLGACTKPGGLMIVIVFCKFGNLSLYLRS 930  
Db 864 KEGATHSEHRALMSELKILIHGHHLNVNLLGACTKPGGLMIVIVFCKFGNLSLYLRS 923  
Qy 931 KRNEFV-----PYTKK--CARFQGGDYVGAIPVDLKRDLSDITSQSSASSGFFVE 979  
Db 924 KRLDFLNDKDALHMEPKKEMPGLEQK-----KPRLDSTVTSSESFASGFGQE 973  
Qy 980 EKSLSDVEEAPEDLYKDFLTHLEHLYCYSFQVAKGMEFLASRKCIRHDLAARNILLSEK 1039  
Db 974 DKSLSDVEEEDSDGFKPEPITWEDLISYFQVARGMEFLSSRKCIRHDLAARNILLSEN 1033  
Qy 1040 NVVKICDFGLARDIYKDPYVRKGDARLPLKWNAPETIEDRVVTIQSDVWSFGVLLWEIF 1099









Db 17 PSLIVCQLLPSILPNEKIVPLSSSFLRCFGESEVSWQHPMSE-EDPNVEIRTEEN 75  
Qy 80 SDGLFCKTLTIPIKVIKNDTCAYKCFYRETDLASV-----IYVVQDYRPFPI-ASVSDQ 132  
Db 76 NSSLFVTVLEVNVNAAHTGWITCYNHTQTESEIEGRHIYIYVDPDMAFVPLGMDTS 135  
Qy 133 HGVVYITENKNTVVIPLCLASISNLVSLCARYPEKRFVDPGNRISWDSKKGFIPSYMI 192  
Db 136 LVIV-----EEDDSAILPCLTDPDTEVTL---HNNGLVLP---ASVDSRQGFN-GTFVS 183  
Qy 193 SYAGMVFCEAKINDESQSIYIVVVVGYRIYDVVLSPSHGIEL-----SVGEKLVN 245  
Db 184 ---GPYICEATVRGTFKT-----SEFNVYALKATSELNLEMDTRQTYKAGETIVVT 233  
Qy 246 CTARTELNVGIDFNWVEPSSKHQHKLVNRDLTKQSGSEMKKFLSTLIDGVTNRSDGLY 305  
Db 234 CAVFN--NEVVDLQWTVGEVRNKGITMLBEIKLPS-----IKLVYTLTVPKATVKSQGDY 287  
Qy 306 TCAASSGLMTKK--NSTFVRVHEKPFVAFGSGMESLVEATVGBRVRIIPAKYLGYPPEIK 363  
Db 288 ECAARQATKEVKEMKTVTISVHEKGFVQIRPTFGHLETVNLHQVREFVVEVQAYPTPRIS 347  
Qy 364 WYKNGIPLESNHTKAGHV-----LTMEVSERTGNYTVILTNPISEKQSHV 412  
Db 348 WLKDNLTLENLTEITTDVORSQETRYQSKLKLIRAKEEDSGHYTIIVQN--DDDMKSYT 405  
Qy 413 VSLVVVYVPOIGESKLSIPVDSYQYGTOTLTCTTVYAIPPPHILHHWQLEBECANPSQ 472  
Db 406 FELSTLPASILE--LVD--DHHGSGGGQTVRCATAGTPIPN----- 443  
Qy 473 AVSVTNPYPCBEVRSDFOGKNKIEVNKNQFALLIEGKNKTSTVLVQAANVSALYKCEA 532  
Db 444 -----IEMWICKDIK-----KC-- 455  
Qy 533 VNKGERSVISEHVTRGPBITLQPMQPTQESVSLWCTADRTSTPENLTWYKLGPPPLP 592  
Db 456 ----- 455  
Qy 593 IHVGELPTPVCKNLTILKLNATMFSNSTNDILIMELKNASLQDQGDYVCLAQDRKTKR 652  
Db 456 -----NNDTSWTVLASVSN-----IITEFH----- 476  
Qy 653 HCVRQLTVLERVAPTITGNLENTTISGESIEVSCVTASGNPPPPQIMWFKDNETLVEDSG 712  
Db 477 -----QRGESVTEGRV--SFAKVEETIAVRCLAK-----NDLG 507  
Qy 713 IVLKQGNRNTIRVRKEDSGLYTCQACSVLGCACKEAFAFIIEGAOEKTNLEIILVGT 772  
Db 508 I-----GNRELKLVAPLSRSE-LTVAAAVLVL-----LVIVIVSL 541  
Qy 773 VIAMFFLLVILIRTVKRANGELK-----TCYLSIVMDPDELPLDEHCERLPPYD 823  
Db 542 IVLVIVW-----KQPRYEIRWRVIESIPDGHEHYIYDPMQLPYD----- 582  
Qy 824 ASKWEPRDLKGLGRGAFGOVTEADAFGIDKATCTAVKMLKEGATHSEHRLM 883  
Db 583 -SRWEPRDLGLVIRILGSGAPKVEGTAYLSRQPMVKVAVKMLKPTARSSEKQALM 641  
Qy 884 SELKILIHGHILNVNLLGACTKPGGLPMVIVEFCFKNLSTYLSRKRNEFV---PYKT 940  
Db 642 SELKIMTHLPHLNIIVNLGACTK-SGPIVIITEYCFYGLVNLVHKNRDSFMSRHEPK 700  
Qy 941 KG-----AFRQGDYVGAIPVDLKRRLDSIYSSQSSASSGF----- 977  
Db 701 KCOLDIFGLNPADESTSYVILSFENNQDYVDMKQADTTQYVPMLEKKEYSKYSQISL 760  
Qy 978 -----VEEKSLSDVEEAPEDLYKDFLLEHLICYSFOVAKGMEFLASRKCITHRDLLA 1031  
Db 761 YDRPASVKKKSLMDSEAKNLLSDDSEGLTLLDLSTFYVQVARGMEFLASKNCVHRDLA 820  
Qy 1032 RNILLSEKNVVKICDFGLARDIYKDDPYVRKGDARLPLKMAPETIFDRVYTTQSDVWSF 1091  
Db 821 RNVLLAQGKIVKICDFGLARDIMHDSNVYSKGTFTPLPVKMAPESIFDNLTYTLLSDVWSY 880

Qy 1092 GVLLWEIETSI GASPYQCVKIDEEFCRLKEGTRMRAPDYTTTPMYQOTMLDCWHGERSORP 1151  
Db 881 GVLLWEIETSI LGGTPIPGMWDSTFYNKIKSGYRMAKPDHATSEYIEIMVQCWNSPEKRP 940  
Qy 1152 TFSLEVEHLGNLLOANAQDGKDIYIIVLPISETLSMEEDSGLSLPTSPVSCMBEVEVCDPK 1211  
Db 941 SPVHLSIEIVENLPG--QYKKS-----EK 963  
Qy 1212 FHYDNTAGISQYLOLNRKRSRPSVKTPEIDPLEEPEV-KVIPDDNQDTSQMWLASBELK 1270  
Db 964 IHLN-----FLKS-----DHPAVARMRVDSNAYIGVTVKNEEDK 998  
Qy 1271 TLE-----DRTKLSPSFGMVPSKRESVASE-----GSNQTSYGSGYHSDDTDTT 1317  
Db 999 LKWEGLDEORLSADSGYIIPDPIDPVEEDLGRNRHSSQTS-ERSAIEYSGSSST 1057  
Qy 1318 VYSSEBAELKLKLEIGVQGTSTAQILQPPDS 1347  
Db 1058 -FKREDETIEDIMDDIGIDSSDLVEDS 1086

## RESULT 13

S33727

platelet-derived growth factor receptor alpha precursor - mouse

N;Contains: protein-tyrosine kinase (EC 2.7.1.112)

C;Species: Mus musculus (house mouse)

C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 05-Oct-2004

C;Accession: I57511; S33727

R;Stiles, C.D.; Wang, C.

Mol. Cell. Biol. 10, 6781-6784, 1990

A;Title: Retinoic acid promotes transcription of the platelet-derived growth factor alpha

A;Reference number: I57511; MUID:91061789; PMID:2174116

A;Accession: I57511

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-1089 &lt;RES&gt;

A;Cross-references: UNIPROT:P26618; UNIPARC:UPI000016CF20; GB:M57683; NID:G199783; PIDN:1

A;Genetics:

A;Gene: PDGF-alpha-R

C;Superfamily: Tyrosine-protein kinase, CSF-1/PDGF receptor type; immunoglobulin homology

C;Keywords: ATP; glycoprotein; growth factor receptor; phosphotransferase; transmembrane

F;1-23/Domain: signal sequence #status predicted &lt;SIG&gt;

F;228-292/Domain: immunoglobulin homology &lt;IMM&gt;

F;591-957/Domain: protein kinase homology &lt;KIN&gt;

F;599-607/Region: protein kinase ATP-binding motif

F;42,76,89,103,179,353,459,468,506/Binding site: carbohydrate (Asn) (covalent) #statu

Query Match 16.7%; Score 1184; DB 1; Length 1089;

Best Local Similarity 26.2%; Pred. No. 6.1e-42;

Matches 371; Conservative 191; Mismatches 417; Indels 436; Gaps 49;

Qy 23 GLPSVSLDPLRLSIQKDLITIKANTLTQITCRQDLDMWPNQSGSORVEV--TECS 80

Db 19 GLISQCLLPILPNEKIVQLNSSFSLRCVGESEVSWQHPMSE-EDPNVEIRSENN 77

Qy 81 DGLFCKTLIPIKVIKNDTCAYKCFYR--ETDLASV-----IYVVQDYRPFPI-ASVSDQH 133

Db 78 SGLFVTVLEVNVNAAHTGWITCYNHTQTESEIEGRHIYIYVDPDMAFVPLGMDTSL 137

Qy 134 GVYVITENKNTVVIPLCLASISNLVSLCARYPEKRFVDPGNRISWDSKKGFIPSYMIS 193

Db 138 VIV-----EEDDSAILPCLTDPDTEVTL---HNNGLVLP---ASVDSRQGFN-GTFVS- 184

Qy 194 YAGMVFCEAKINDESQSIYIVVVVGYRIYDVVLSPSHGIEL-----SVGEKLVNLC 246

Db 185 --GPYICEAAVKGRTFKT-----SAFNVYALKATSELNLEMDARQTVVYKAGETIVVTC 235

Qy 247 TARTELNVGIDFNWVEPSSKHQHKLVNRDLTKQSGSEMKKFLSTLIDGVTNRSDGLYT 306

Db 236 AVFN--NEVVDLQWTVYGGVRNKGITMLBEIKLPS-----IKVYTLTVPKATVKDSGEYE 289

Qy 307 CAASSGLMTKK--NSTFVRVHEKPFVAFGSGMESLVEATVGBRVRIIPAKYLGYPPEIKW 364









```

Db      768  QGLDFLASRNCIHRDVARNVLLTDKRVAKICDFGLARDIMNDSNYVVKGNARLPVKWMA 827
Qy      1074 PETIFDRVYTIQSDVWSFGVLLWEIFSLGASPYGVKIDEEFCRRLKKGTRMRAPDYTP 1133
Db      828  PESIFDCVYTVQSDVWSYGILLWEIFSLGKSPYPSMAVDSRFYKMKVKRGYQMSQDPFALP 887
Qy      1134 EMYCTMLDCWHGEPSPRPTSELSVEHLGNLLQANAQODGKDYIVLPISETLSMEEDSGLS 1193
Db      888  EITYMIMKMCWNLBPTERTFTSMISQMINRLLGGQDOEKLIY-----RNVQPEQ----- 936
Qy      1194 LPTSPVSCMEEEVEVCDPKPHYD 1215
Db      937  -----VAEGEACDEPKRYD 950

```

Search completed: March 10, 2006, 18:38:07  
 Job time : 44.0732 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2006, 18:36:34 ; Search time 33.7622 Seconds  
(without alignments)  
3320.525 Million cell updates/sec

Title: US-10-633-742-6

Perfect score: 7095

Sequence: 1 MESKVLAVLWLCVETRAA.....GSTAQILQPSGTTLSPPV 1356

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.\*  
1: /cgm2\_6/ptodata/1/iaa/5\_COMB.pep.\*  
2: /cgm2\_6/ptodata/1/iaa/6\_COMB.pep.\*  
3: /cgm2\_6/ptodata/1/iaa/H\_COMB.pep.\*  
4: /cgm2\_6/ptodata/1/iaa/PTUS\_COMB.pep.\*  
5: /cgm2\_6/ptodata/1/iaa/RE\_COMB.pep.\*  
6: /cgm2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7095	100.0	1356	2	US-09-098-707A-2
2	7095	100.0	1356	2	US-09-483-539-2
3	7095	100.0	1356	2	US-10-100-405A-2
4	7095	100.0	1356	2	US-10-022-939-2
5	7092	99.9	1356	2	US-09-949-016-6198
6	7092	99.9	1456	2	US-09-949-016-9853
7	7070	99.6	1356	1	US-08-810-116-8
8	7070	99.6	1356	1	US-07-930-548A-8
9	6106.5	86.1	1367	1	US-08-443-861-2
10	6106.5	86.1	1367	2	US-08-193-829B-2
11	6106.5	86.1	1367	2	US-09-766-678-2
12	6097.5	85.9	1367	1	US-07-813-593-4
13	6097.5	85.9	1367	1	US-07-977-451-6
14	6097.5	85.9	1367	1	US-07-946-507-4
15	6097.5	85.9	1367	1	US-08-252-517-6
16	6097.5	85.9	1367	1	US-07-906-397A-6
17	6097.5	85.9	1367	1	US-08-601-891-6
18	6097.5	85.9	1367	1	US-09-021-324-6
19	6097.5	85.9	1367	2	US-09-872-136B-6
20	6097.5	85.9	1367	2	US-09-919-408A-6
21	6097.5	85.9	1367	4	PCT-US92-02750-8
22	6097.5	85.9	1367	4	PCT-US92-05401-6
23	6097.5	85.9	1367	4	PCT-US92-09893-6
24	4101.5	57.8	788	1	US-08-232-538-15
25	4101.5	57.8	788	1	US-08-786-164-15
26	4020	56.7	764	2	US-09-142-956B-14
27	4017	56.6	767	1	US-08-874-678-2

28 4017 56.6 767 2 US-08-643-839-2 Sequence 2, Appli  
29 4017 56.6 767 2 US-09-348-886-2 Sequence 2, Appli  
30 4017 56.6 767 2 US-10-105-901A-2 Sequence 2, Appli  
31 3491 49.2 668 1 US-08-232-538-13 Sequence 13, Appl  
32 3491 49.2 668 1 US-08-786-164-13 Sequence 13, Appl  
33 3453 48.7 806 1 US-08-443-861-5 Sequence 5, Appli  
34 3453 48.7 806 1 US-08-193-829B-5 Sequence 5, Appli  
35 3453 48.7 806 2 US-09-766-678-5 Sequence 5, Appli  
36 3437 48.4 805 2 US-08-985-526-34 Sequence 19, Appl  
37 2804 39.5 1363 2 US-09-375-248-19 Sequence 19, Appl  
38 2797.5 39.4 1368 1 US-08-643-839-34 Sequence 34, Appl  
39 2797.5 39.4 1368 2 US-09-348-886-34 Sequence 34, Appl  
40 2797.5 39.4 1368 2 US-10-105-901A-34 Sequence 34, Appl  
41 2797.5 39.4 1368 2 US-08-340-011-4 Sequence 4, Appli  
42 2759 38.9 1363 1 US-08-901-710-4 Sequence 4, Appli  
43 2759 38.9 1363 2 US-09-375-248-2 Sequence 2, Appli  
44 2759 38.9 1363 2 US-09-169-079-4 Sequence 4, Appli  
45 2759 38.9 1363 2

## ALIGNMENTS

### RESULT 1

US-09-098-707A-2  
; Sequence 2, Application US/09098707A  
; Patent No. 6204011  
; GENERAL INFORMATION:  
; APPLICANT: Kendall, Richard L.  
; Thomas, Kenneth A.  
; Mao, Xianzhi  
; Tebben, Andrew J.  
; TITLE OF INVENTION: HUMAN RECEPTOR TYROSINE KINASE, KDR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marck & Co., Inc.  
; STREET: P.O. Box 2000  
; CITY: Rahway  
; STATE: NJ  
; COUNTRY: US  
; ZIP: 07065-0907  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/098,707A  
; FILING DATE: 17-Jun-1998  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hand, J. Mark  
; REGISTRATION NUMBER: 36,545  
; REFERENCE/DOCKET NUMBER: 19963PV  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 732/594-3905  
; TELEFAX: 732/594-4720  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1356 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-098-707A-2

Query Match 100.0%; Score 7095; DB 2; Length 1356;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESKVLAVLWLCVETRAASVGLPSVSLDPRLSIQDKILTIKANTTLOITCRGQDLD 60  
Db 1 MESKVLAVLWLCVETRAASVGLPSVSLDPRLSIQDKILTIKANTTLOITCRGQDLD 60

QY 61 WLFPNNSGSEQRVEVTECSGDLGFKTLTIPKVIKNDGTGAYKCFYRETDLASVIVYVQD 120  
DB 61 WLFPNNSGSEQRVEVTECSGDLGFKTLTIPKVIKNDGTGAYKCFYRETDLASVIVYVQD 120  
QY 121 YRSPFIASVSDQHGVVITENKNTVVIKLGSIISNLNLSLCARYPKRFPVPGNRIISWD 180  
DB 121 YRSPFIASVSDQHGVVITENKNTVVIKLGSIISNLNLSLCARYPKRFPVPGNRIISWD 180  
QY 181 SKGFTTIPSYMISYAGWVFCEAKINDESYSIMYIVVVGVRIYDVVLSPSHGLEISVGE 240  
DB 181 SKGFTTIPSYMISYAGWVFCEAKINDESYSIMYIVVVGVRIYDVVLSPSHGLEISVGE 240  
QY 241 KVLNCTARTLNKLVGIDFNWEYSSKHQKHLNVRDLKTQSGEMKFLSTLTIDGVTRS 300  
DB 241 KVLNCTARTLNKLVGIDFNWEYSSKHQKHLNVRDLKTQSGEMKFLSTLTIDGVTRS 300  
QY 301 DOGLYTCAASGLMTKNTSTFVRVHEKPFVAFSGMESLVEATVGERVRIPAKYLGYPVP 360  
DB 301 DOGLYTCAASGLMTKNTSTFVRVHEKPFVAFSGMESLVEATVGERVRIPAKYLGYPVP 360  
QY 361 EIKWYKNGIPLSNHTTKAGHVLTIMEVSRDGTGNTVILTNPISKEKQSHVSVLVVYVP 420  
DB 361 EIKWYKNGIPLSNHTTKAGHVLTIMEVSRDGTGNTVILTNPISKEKQSHVSVLVVYVP 420  
QY 421 POIGEXSLIPVDSYQYGTTLTCTVYALPPPHHHIHWYQLEBEECANEPSQAVSVTNPY 480  
DB 421 POIGEXSLIPVDSYQYGTTLTCTVYALPPPHHHIHWYQLEBEECANEPSQAVSVTNPY 480  
QY 481 PCBEWRSVEDFQGGNKTIEVKNQFALJEGNKTSTLVIQAANVSALYKCEAVNKVGRGE 540  
DB 481 PCBEWRSVEDFQGGNKTIEVKNQFALJEGNKTSTLVIQAANVSALYKCEAVNKVGRGE 540  
QY 541 RVISFHVTRGPETLQDPMQPTQESVSLACTADRSTFENLTWYKLGQPLPHVGELEPT 600  
DB 541 RVISFHVTRGPETLQDPMQPTQESVSLACTADRSTFENLTWYKLGQPLPHVGELEPT 600  
QY 601 PVCKNLDLTKLKNATMESNTDILIMELKNASLQDQGVVCLAQDRKTKKRHCVVRLT 660  
DB 601 PVCKNLDLTKLKNATMESNTDILIMELKNASLQDQGVVCLAQDRKTKKRHCVVRLT 660  
QY 661 VLERVAPITGNLENQTTSGESIEVSCTAGNPPQIMMFKNETLVEDSGIVLKDGNR 720  
DB 661 VLERVAPITGNLENQTTSGESIEVSCTAGNPPQIMMFKNETLVEDSGIVLKDGNR 720  
QY 721 NLTIIRVRKEDEGLYTQACSVLGCACVFAFFIIEGAQKTNLEIILVGTAVIAMFFWL 780  
DB 721 NLTIIRVRKEDEGLYTQACSVLGCACVFAFFIIEGAQKTNLEIILVGTAVIAMFFWL 780  
QY 781 LLVILRTVVRANGGELKTGYLSIVMDPDELPLDEHCERLPYDASKWEPFRDLKLGKPL 840  
DB 781 LLVILRTVVRANGGELKTGYLSIVMDPDELPLDEHCERLPYDASKWEPFRDLKLGKPL 840  
QY 841 GRGAFQVIBADAFIDKATCTRTAVKMLKEGATHSEHRLMSELKILIHGHILNVN 900  
DB 841 GRGAFQVIBADAFIDKATCTRTAVKMLKEGATHSEHRLMSELKILIHGHILNVN 900  
QY 901 LLGACTKPGGPLMVIIEFCFKGNLSTVLSKRNEFPYKTKGARPRQGDYVCAIPVDLK 960  
DB 901 LLGACTKPGGPLMVIIEFCFKGNLSTVLSKRNEFPYKTKGARPRQGDYVCAIPVDLK 960  
QY 961 RRLDSITSSQSSASSGVEBSKLSDBEEBAPDLVKDPLTLEHLICYSPQVAKGMEFLA 1020  
DB 961 RRLDSITSSQSSASSGVEBSKLSDBEEBAPDLVKDPLTLEHLICYSPQVAKGMEFLA 1020  
QY 1021 SRKCIHRDLAARNILISEKNVVKICDFGLARDIYKOPDYVRKGDARLPLKWWAPETIFDR 1080  
DB 1021 SRKCIHRDLAARNILISEKNVVKICDFGLARDIYKOPDYVRKGDARLPLKWWAPETIFDR 1080  
QY 1081 VYTIQSDWRSFVLLWEIFSLGASPYGVKIDBEFCRLKEGTRMRAPDYTTPEMYQTML 1140  
DB 1081 VYTIQSDWRSFVLLWEIFSLGASPYGVKIDBEFCRLKEGTRMRAPDYTTPEMYQTML 1140

QY 1141 DCWHGSPSQRTFSELVEHLGNLQANAQDGKDYIVLPISETLSMEEDSGLSPTSPVS 1200  
DB 1141 DCWHGSPSQRTFSELVEHLGNLQANAQDGKDYIVLPISETLSMEEDSGLSPTSPVS 1200  
QY 1201 CMEEBVCDPKFHYDNTAGISQYQLNSKRSRPSVKTTFEDIPLEBPEVKVIPPDDNQDTS 1260  
DB 1201 CMEEBVCDPKFHYDNTAGISQYQLNSKRSRPSVKTTFEDIPLEBPEVKVIPPDDNQDTS 1260  
QY 1261 GMYLASEELKTLDRDKLSPSGMGVPSKRSRESVASEGNSQTSQYSGYHSDDDTDTTVYS 1320  
DB 1261 GMYLASEELKTLDRDKLSPSGMGVPSKRSRESVASEGNSQTSQYSGYHSDDDTDTTVYS 1320  
QY 1321 SEEAELKLEIGVQGTGTAQILQPDSTGTTLSPPV 1356  
DB 1321 SEEAELKLEIGVQGTGTAQILQPDSTGTTLSPPV 1356

## RESULT 2

US-09-483-539-2  
; Sequence 2, Application US/09483539  
; Patent No. 6359115  
; GENERAL INFORMATION:  
; APPLICANT: Kendall, Richard L.  
; APPLICANT: Thomas, Kenneth A.  
; APPLICANT: Mac, Xianzhi  
; APPLICANT: Tebben, Andrew J.  
; TITLE OF INVENTION: HUMAN RECEPTOR TYROSINE KINASE, KDR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merck & Co., Inc.  
; STREET: P.O. Box 2000  
; CITY: Rahway  
; STATE: NJ  
; COUNTRY: US  
; ZIP: 07065-0907  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/483,539  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hand, J. Mark  
; REGISTRATION NUMBER: 36,545  
; REFERENCE/DOCKET NUMBER: 19963PV  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 732/594-3905  
; TELEFAX: 732/594-4720  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1356 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-483-539-2

Query Match 100.0%; Score 7095; DB 2; Length 1356;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESKVLALVALMVCVETRAASVGLPSVSLDLRLSLSTOKDILITKANTTLOITCRGORDLD 60  
DB 1 MESKVLALVALMVCVETRAASVGLPSVSLDLRLSLSTOKDILITKANTTLOITCRGORDLD 60  
QY 61 WLFPNNSGSEQRVEVTECSGDLGFKTLTIPKVIKNDGTGAYKCFYRETDLASVIVYVQD 120  
DB 61 WLFPNNSGSEQRVEVTECSGDLGFKTLTIPKVIKNDGTGAYKCFYRETDLASVIVYVQD 120  
QY 121 YRSPFIASVSDQHGVVITENKNTVVIKLGSIISNLNLSLCARYPKRFPVPGNRIISWD 180  
DB 121 YRSPFIASVSDQHGVVITENKNTVVIKLGSIISNLNLSLCARYPKRFPVPGNRIISWD 180

Db 121 YRSPFIASVSDQHGIVVYITENKNTVIPCGLSGISNLNLSLCARYPEKRFVDPDGNRISWD 180  
Qy 181 SKGFTIPSYMIYAGWVFCEAKINDESYQIMIVVVGRIYDVVLSPHSGHIELSVGE 240  
Db 181 SKGFTIPSYMIYAGWVFCEAKINDESYQIMIVVVGRIYDVVLSPHSGHIELSVGE 240  
Qy 241 KLVLNCTARTLNAGIDFNWYPSGKHQKLVNRDLKTQSGSEMKFSLTLTIDGVTRS 300  
Db 241 KLVLNCTARTLNAGIDFNWYPSGKHQKLVNRDLKTQSGSEMKFSLTLTIDGVTRS 300  
Qy 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFGSGMESLVEATVGERVIRIPAKYLGYP 360  
Db 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFGSGMESLVEATVGERVIRIPAKYLGYP 360  
Qy 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTNYTILNTPISKQSHVSVLVYVP 420  
Db 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTNYTILNTPISKQSHVSVLVYVP 420  
Qy 421 POIGKSLIPVDSYQVGTOTLTCTVYAI PPPHHIHWYQLEBECANEPQSVVTNYP 480  
Db 421 POIGKSLIPVDSYQVGTOTLTCTVYAI PPPHHIHWYQLEBECANEPQSVVTNYP 480  
Qy 481 PCEWRSVEDFQGNKIEVKNQFALIBGNKTVSTLVIQANVSALYKCEAVNKVGRGE 540  
Db 481 PCEWRSVEDFQGNKIEVKNQFALIBGNKTVSTLVIQANVSALYKCEAVNKVGRGE 540  
Qy 541 RVISFHVTRGPEITLQPMQPTQESVSLWCTADRSTFENLTWYKLGQPPIHVGELPT 600  
Db 541 RVISFHVTRGPEITLQPMQPTQESVSLWCTADRSTFENLTWYKLGQPPIHVGELPT 600  
Qy 601 PCKNLDLTKLNATMFNSNDILIMELKNASIQDQDYVCLAQDRKTKKHCHCVRLT 660  
Db 601 PCKNLDLTKLNATMFNSNDILIMELKNASIQDQDYVCLAQDRKTKKHCHCVRLT 660  
Qy 661 VLVRVAPITCNLENQTTISGESIEVSTASGNPPQIMWPKNETLVEDSGIVLKDGNR 720  
Db 661 VLVRVAPITCNLENQTTISGESIEVSTASGNPPQIMWPKNETLVEDSGIVLKDGNR 720  
Qy 721 NLATRRVRKDEGLYTQACSVLGCACKVEAFPIIEGAQKTNLBIILVGTAVIAMEFWL 780  
Db 721 NLATRRVRKDEGLYTQACSVLGCACKVEAFPIIEGAQKTNLBIILVGTAVIAMEFWL 780  
Qy 781 LVLVILRTVKRANGELKTGYSIVMDPELPLDEHCERLPYDASKWEFFPRDLKLGKPL 840  
Db 781 LVLVILRTVKRANGELKTGYSIVMDPELPLDEHCERLPYDASKWEFFPRDLKLGKPL 840  
Qy 841 GRGAFQGVIEADAFGIDKATCRTVAVKMLKEGATHSEHRAIMSELKILIHGHILNVN 900  
Db 841 GRGAFQGVIEADAFGIDKATCRTVAVKMLKEGATHSEHRAIMSELKILIHGHILNVN 900  
Qy 901 LLGACTKPGGLMWIVFCKFNGNLSTYLRSKRNEFPVYKTKGARFRQGDYVGAIPVDLK 960  
Db 901 LLGACTKPGGLMWIVFCKFNGNLSTYLRSKRNEFPVYKTKGARFRQGDYVGAIPVDLK 960  
Qy 961 RRLDSITSSQSSASSGFVEEKSLSDBVEEBEAPEDLYKDFLTLHLICYSFOVAKGMEFLA 1020  
Db 961 RRLDSITSSQSSASSGFVEEKSLSDBVEEBEAPEDLYKDFLTLHLICYSFOVAKGMEFLA 1020  
Qy 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDDYVRKGDARLPKWMAPETIFDR 1080  
Db 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDDYVRKGDARLPKWMAPETIFDR 1080  
Qy 1081 VYTIQSDWMSFGVLLWEIFSLGASPYPGVKIDEEFCRLKEGTRMRAPDYTPPMYQTM 1140  
Db 1081 VYTIQSDWMSFGVLLWEIFSLGASPYPGVKIDEEFCRLKEGTRMRAPDYTPPMYQTM 1140  
Qy 1141 DCWHGERSQRTFSELVEHLGNLQANAQDGKDYIVLPISSETLSMEEDSGLSLPTSPVS 1200  
Db 1141 DCWHGERSQRTFSELVEHLGNLQANAQDGKDYIVLPISSETLSMEEDSGLSLPTSPVS 1200  
Qy 1201 CMEEBEECDPKFHYDNTAGISQYLONSKRKRSPVSKTFEDIPLEEBEVKVIPODNDQDS 1260  
Db 1201 CMEEBEECDPKFHYDNTAGISQYLONSKRKRSPVSKTFEDIPLEEBEVKVIPODNDQDS 1260

Qy 1261 GMLASBELKLTLEDRTKLSPSFGMWPSKRESVASSEGSNOTSGYQSGYHSDDDTDTTVYS 1320  
Db 1261 GMLASBELKLTLEDRTKLSPSFGMWPSKRESVASSEGSNOTSGYQSGYHSDDDTDTTVYS 1320  
Qy 1321 SEEAELKLKLEIGVQGTSTAQILOPDSGTTILSSPPV 1356  
Db 1321 SEEAELKLKLEIGVQGTSTAQILOPDSGTTILSSPPV 1356

## RESULT 3

US-10-100-405A-2  
; Sequence 2, Application US/10100405A  
; Patent No. 6841367  
; GENERAL INFORMATION:  
; APPLICANT: Kendall, Richard L.  
; APPLICANT: Thomas, Kenneth A.  
; APPLICANT: Mac, Xianzhi  
; APPLICANT: Tebben, Andrew  
; TITLE OF INVENTION: HUMAN RECEPTOR TYROSINE KINASE, KDR  
; FILE REFERENCE: 19963YDC  
; CURRENT APPLICATION NUMBER: US/10100.405A  
; CURRENT FILING DATE: 2002-08-13  
; PRIOR APPLICATION NUMBER: 10/022,939  
; PRIOR FILING DATE: 2001-12-18  
; PRIOR APPLICATION NUMBER: 09/483,539  
; PRIOR FILING DATE: 2000-01-14  
; PRIOR APPLICATION NUMBER: 09/098,707  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/050,962  
; PRIOR FILING DATE: 1997-06-18  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; TYPE: PRT  
; ORGANISM: Human  
US-10-100-405A-2

Query Match 100.0%; Score 7095; DB 2; Length 1356;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESKVLAVLWLCVETRAASVGLPSVSLDPLRISIQDKILTIKANTLQITCGQRDL 60  
Db 1 MESKVLAVLWLCVETRAASVGLPSVSLDPLRISIQDKILTIKANTLQITCGQRDL 60  
Qy 61 WLPNNQSGSEQRVEVTECSGDLFCCKTLTI PKVIGNDTGAYKCFYRETDLASVIYVQD 120  
Db 61 WLPNNQSGSEQRVEVTECSGDLFCCKTLTI PKVIGNDTGAYKCFYRETDLASVIYVQD 120  
Qy 121 YRSPFIASVSDQHGIVVYITENKNTVIPCGLSGISNLNLSLCARYPEKRFVDPDGNRISWD 180  
Db 121 YRSPFIASVSDQHGIVVYITENKNTVIPCGLSGISNLNLSLCARYPEKRFVDPDGNRISWD 180  
Qy 181 SKGFTIPSYMIYAGWVFCEAKINDESYQIMIVVVGRIYDVVLSPHSGHIELSVGE 240  
Db 181 SKGFTIPSYMIYAGWVFCEAKINDESYQIMIVVVGRIYDVVLSPHSGHIELSVGE 240  
Qy 241 KLVLNCTARTLNAGIDFNWYPSGKHQKLVNRDLKTQSGSEMKFSLTLTIDGVTRS 300  
Db 241 KLVLNCTARTLNAGIDFNWYPSGKHQKLVNRDLKTQSGSEMKFSLTLTIDGVTRS 300  
Qy 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFGSGMESLVEATVGERVIRIPAKYLGYP 360  
Db 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFGSGMESLVEATVGERVIRIPAKYLGYP 360  
Qy 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTNYTILNTPISKQSHVSVLVYVP 420  
Db 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTNYTILNTPISKQSHVSVLVYVP 420  
Qy 421 POIGKSLIPVDSYQVGTOTLTCTVYAI PPPHHIHWYQLEBECANEPQSVVTNYP 480

```
Db 421 PQIGKSLISPVDYQYGTQTTLCTVYAI PPPHHIHWYQLEBECANEPQAVSVTNPY 480
Qy 481 PCEWRSVEDFQGNKIEVKNQFALIEGKNTVSTLVIQAAVNSALYKCEAVNKVGRGE 540
Db 481 PCEWRSVEDFQGNKIEVKNQFALIEGKNTVSTLVIQAAVNSALYKCEAVNKVGRGE 540
Qy 541 RVISFHVTRGPEITLPDMQPTQESVSLMCTADRSTFENLTWYKLGQPLPIHVGLPT 600
Db 541 RVISFHVTRGPEITLPDMQPTQESVSLMCTADRSTFENLTWYKLGQPLPIHVGLPT 600
Qy 601 PVCKNLDLTKLNATMFSNSTNDILIMELKNASLQDQGVYVCLAQDRKTKRHCVRQLT 660
Db 601 PVCKNLDLTKLNATMFSNSTNDILIMELKNASLQDQGVYVCLAQDRKTKRHCVRQLT 660
Qy 661 VLERVAPITITGNLENQTTISGESIEVSCASGNPPQIMWFKONETLVEDSGIVLKDGNR 720
Db 661 VLERVAPITITGNLENQTTISGESIEVSCASGNPPQIMWFKONETLVEDSGIVLKDGNR 720
Qy 721 NLTIIRVRKEDGLYTQACSVLGCACVAFPIIEGAQEKTNLEIILVGTAVIAMFFWL 780
Db 721 NLTIIRVRKEDGLYTQACSVLGCACVAFPIIEGAQEKTNLEIILVGTAVIAMFFWL 780
Qy 781 LLVILITVRANGELKGTGLSVLMDPDELPLDEHCERLPYDASKWEFFPRDRILKLGKPL 840
Db 781 LLVILITVRANGELKGTGLSVLMDPDELPLDEHCERLPYDASKWEFFPRDRILKLGKPL 840
Qy 841 GRGAFGOVIBADAFIDKATCRTVAVKMLKEGATHSEHRLMSLKIILIHGHILNVN 900
Db 841 GRGAFGOVIBADAFIDKATCRTVAVKMLKEGATHSEHRLMSLKIILIHGHILNVN 900
Qy 901 LLGACTPGGGLMVIIVFCKFNGNLSTYLSRKNEFVYKTKGARFRQGDYVGAIPVDLK 960
Db 901 LLGACTPGGGLMVIIVFCKFNGNLSTYLSRKNEFVYKTKGARFRQGDYVGAIPVDLK 960
Qy 961 RLDSITSSQSSASSGFVBEKSLSDVEEBEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020
Db 961 RLDSITSSQSSASSGFVBEKSLSDVEEBEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020
Qy 1021 SRKCIHRDLAARNLILSEKNVVKICDFGLARDYKDPDYVRKGDARLPLKWMAPETIFDR 1080
Db 1021 SRKCIHRDLAARNLILSEKNVVKICDFGLARDYKDPDYVRKGDARLPLKWMAPETIFDR 1080
Qy 1081 VYTIQSDVWSFGVLLWEIFSLGASPYGVKIDBEFCRRLKEGTRMRAPDYTTPEMYQTML 1140
Db 1081 VYTIQSDVWSFGVLLWEIFSLGASPYGVKIDBEFCRRLKEGTRMRAPDYTTPEMYQTML 1140
Qy 1141 DWHGEPSPRPTSELVEHNLGNLQAAQDGKDYIVLPISETLSMEEDSGLSLPTSPVS 1200
Db 1141 DWHGEPSPRPTSELVEHNLGNLQAAQDGKDYIVLPISETLSMEEDSGLSLPTSPVS 1200
Qy 1201 CMEEEEVCDPKFHYDNTAGISQYLQNSKRKSRPVSVKTPEDIPLERPEVKVIPPDDNQDTS 1260
Db 1201 CMEEEEVCDPKFHYDNTAGISQYLQNSKRKSRPVSVKTPEDIPLERPEVKVIPPDDNQDTS 1260
Qy 1261 GMVLASEELKLEDRTKLSFGWMPKSRVSESGNSQTSQYSGYHSDDTDTTVYS 1320
Db 1261 GMVLASEELKLEDRTKLSFGWMPKSRVSESGNSQTSQYSGYHSDDTDTTVYS 1320
Qy 1321 SEEAELKLEIGVQTGSTAQIILQPSGTTLSPPV 1356
Db 1321 SEEAELKLEIGVQTGSTAQIILQPSGTTLSPPV 1356
```

## RESULT 4

```
US-10-022-939-2
; Sequence 2, Application US/10022939
; Patent No. 6841382
; GENERAL INFORMATION:
; APPLICANT: Kendall, Richard L.
; APPLICANT: Thomas, Kenneth A.
; APPLICANT: Mao, Xianzhi
; APPLICANT: Tebben, Andrew
; TITLE OF INVENTION: HUMAN RECEPTOR TYROSINE KINASE, KDR
```

```
; FILE REFERENCE: 19963YDB
; CURRENT APPLICATION NUMBER: US/10/022,939
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 09/483,539
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 09/098,707
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/050,962
; PRIOR FILING DATE: 1997-06-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1356
; TYPE: PRT
; ORGANISM: Human
US-10-022-939-2
```

Query Match 100.0%; Score 7095; DB 2; Length 1356;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 MESKVLALVALMVCVETRAASVGLPSVSLDLPRLSTOKDILITIKANTTIQITCRGQDLD 60
Db 1 MESKVLALVALMVCVETRAASVGLPSVSLDLPRLSTOKDILITIKANTTIQITCRGQDLD 60
Qy 61 WLWPNQSGSEQRVEVTECSGDLFCCKTLTIPIKVIKNDTGAYKCFYRETDLASVIVYVQD 120
Db 61 WLWPNQSGSEQRVEVTECSGDLFCCKTLTIPIKVIKNDTGAYKCFYRETDLASVIVYVQD 120
Qy 121 YRSPFIASVSDQHGVVYITENKNTVIVPCLSISNLNLSLCARYPEKRPVDPGNRISWD 180
Db 121 YRSPFIASVSDQHGVVYITENKNTVIVPCLSISNLNLSLCARYPEKRPVDPGNRISWD 180
Qy 181 SKKGFTIPSYMTSYAGWVCEAKINDESQSYIMYIVVVVGYRIYDVVLSHGIELSUGE 240
Db 181 SKKGFTIPSYMTSYAGWVCEAKINDESQSYIMYIVVVVGYRIYDVVLSHGIELSUGE 240
Qy 241 KLVLNCTARTELNVGIDFNWEPSSKHQKLVNRDLKTQSGSEMKKFLSTLTIDGVTRS 300
Db 241 KLVLNCTARTELNVGIDFNWEPSSKHQKLVNRDLKTQSGSEMKKFLSTLTIDGVTRS 300
Qy 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGRVRIPAKYLYGPPP 360
Db 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGRVRIPAKYLYGPPP 360
Qy 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSDRTGNVTIVLTNPISKEKSHVSVLVVYVP 420
Db 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSDRTGNVTIVLTNPISKEKSHVSVLVVYVP 420
Qy 421 PQIGKSLISPVDYQYGTQTTLCTVYAI PPPHHIHWYQLEBECANEPQAVSVTNPY 480
Db 421 PQIGKSLISPVDYQYGTQTTLCTVYAI PPPHHIHWYQLEBECANEPQAVSVTNPY 480
Qy 481 PCEWRSVEDFQGNKIEVKNQFALIEGKNTVSTLVIQAAVNSALYKCEAVNKVGRGE 540
Db 481 PCEWRSVEDFQGNKIEVKNQFALIEGKNTVSTLVIQAAVNSALYKCEAVNKVGRGE 540
Qy 541 RVISFHVTRGPEITLPDMQPTQESVSLMCTADRSTFENLTWYKLGQPLPIHVGLPT 600
Db 541 RVISFHVTRGPEITLPDMQPTQESVSLMCTADRSTFENLTWYKLGQPLPIHVGLPT 600
Qy 601 PVCKNLDLTKLNATMFSNSTNDILIMELKNASLQDQGVYVCLAQDRKTKRHCVRQLT 660
Db 601 PVCKNLDLTKLNATMFSNSTNDILIMELKNASLQDQGVYVCLAQDRKTKRHCVRQLT 660
Qy 661 VLERVAPITITGNLENQTTISGESIEVSCASGNPPQIMWFKONETLVEDSGIVLKDGNR 720
Db 661 VLERVAPITITGNLENQTTISGESIEVSCASGNPPQIMWFKONETLVEDSGIVLKDGNR 720
Qy 721 NLTIIRVRKEDGLYTQACSVLGCACVAFPIIEGAQEKTNLEIILVGTAVIAMFFWL 780
Db 721 NLTIIRVRKEDGLYTQACSVLGCACVAFPIIEGAQEKTNLEIILVGTAVIAMFFWL 780
```

QY 781 LLVILRTVVRANGGELKTYLSIVMDPDELPLDHCERLPYDASKWFFPRDRRLKLGKPL 840  
DB 781 LLVILRTVVRANGGELKTYLSIVMDPDELPLDHCERLPYDASKWFFPRDRRLKLGKPL 840  
QY 841 GRGAFQGVIEADAFIDKATCTRTVAVKMLKEGATHSEHRALMSELKILIHGHLLNVN 900  
DB 841 GRGAFQGVIEADAFIDKATCTRTVAVKMLKEGATHSEHRALMSELKILIHGHLLNVN 900  
QY 901 LLGACTKPGGLMWIVVEFCFKNLSTYLSRKNEFVYKTKGARFRQGDYVGAIPVDLK 960  
DB 901 LLGACTKPGGLMWIVVEFCFKNLSTYLSRKNEFVYKTKGARFRQGDYVGAIPVDLK 960  
QY 961 RRLDSITSSOSSASSGFVEEKSLSDEVEEAPEDLYKDELTLHLICYSFOVAKGMEFLA 1020  
DB 961 RRLDSITSSOSSASSGFVEEKSLSDEVEEAPEDLYKDELTLHLICYSFOVAKGMEFLA 1020  
QY 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPKWMAPETIFDR 1080  
DB 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPKWMAPETIFDR 1080  
QY 1081 VYTIQSDVWFGVLLWEIFSLGASPYGVKIDEEFCRRLKEGTRMRAPDYTTPEMYQTML 1140  
DB 1081 VYTIQSDVWFGVLLWEIFSLGASPYGVKIDEEFCRRLKEGTRMRAPDYTTPEMYQTML 1140  
QY 1141 DCWHGEPQSPTFSELVEHLNLTLOANAQQDKDYIVLPISETLSMEEDSGLSLPTSPVS 1200  
DB 1141 DCWHGEPQSPTFSELVEHLNLTLOANAQQDKDYIVLPISETLSMEEDSGLSLPTSPVS 1200  
QY 1201 CMEEEEVCDPKFHYDNTAGISQYLSQNSKRKRPVSVTKTFEDIPLEBPVKVIPPDDNQDTS 1260  
DB 1201 CMEEEEVCDPKFHYDNTAGISQYLSQNSKRKRPVSVTKTFEDIPLEBPVKVIPPDDNQDTS 1260  
QY 1261 GMLVASELKTLEDRTKLSPSGGMVPSKRSVASEGNSQTSQYSGYHSDDDTTTVYS 1320  
DB 1261 GMLVASELKTLEDRTKLSPSGGMVPSKRSVASEGNSQTSQYSGYHSDDDTTTVYS 1320  
QY 1321 SEAEELKLHIGVQTSQAQILQPSGTTLSPPV 1356  
DB 1321 SEAEELKLHIGVQTSQAQILQPSGTTLSPPV 1356

## RESULT 5

US-09-949-016-6198  
; Sequence 6198, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6198  
; LENGTH: 1356  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-6198

Query Match 99.9%; Score 7092; DB 2; Length 1356;  
Best Local Similarity 99.9%; Pred. NO. 0;  
Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESKVLLAVLWLCVETRAASVGLPSVSLDPLRLSIQKDLTIKANTTLQITCRGQDLD 60  
DB 1 MQSKVLLAVLWLCVETRAASVGLPSVSLDPLRLSIQKDLTIKANTTLQITCRGQDLD 60

QY 61 WLPNPNQSGSEORVEVECTCDGLFCCKTLTIPIKVIGNDTGAYKCFYRETDLASVIVYVQD 120  
DB 61 WLPNPNQSGSEORVEVECTCDGLFCCKTLTIPIKVIGNDTGAYKCFYRETDLASVIVYVQD 120  
QY 121 YRSPFFIASVSDQGVVYITENKNTKTVIICLGSTINLNSLCLARYPEKRPVDPGNRISWD 180  
DB 121 YRSPFFIASVSDQGVVYITENKNTKTVIICLGSTINLNSLCLARYPEKRPVDPGNRISWD 180  
QY 181 SKKGFTIPSMYIYAGMVFCEAKINDESYOSIMYIVVVVGYRIYDVVLSPSHGIELSVGE 240  
DB 181 SKKGFTIPSMYIYAGMVFCEAKINDESYOSIMYIVVVVGYRIYDVVLSPSHGIELSVGE 240  
QY 241 KVLNCTARTALNVGIDFNWEYPSKHQKLVNRDLKTQSGSEMKKFLSLTITDGVTRS 300  
DB 241 KVLNCTARTALNVGIDFNWEYPSKHQKLVNRDLKTQSGSEMKKFLSLTITDGVTRS 300  
QY 301 DOGLYTCAASSGLMTKCNSTFVRVHEKPFVAFSGMESLVEATVGERVRIPAKYLGYPVP 360  
DB 301 DOGLYTCAASSGLMTKCNSTFVRVHEKPFVAFSGMESLVEATVGERVRIPAKYLGYPVP 360  
QY 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDTCNVTYVILTNPISEKQSHVSVLVYVP 420  
DB 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDTCNVTYVILTNPISEKQSHVSVLVYVP 420  
QY 421 PQIGEKSLISPVDSYQYGTTLCTVYAIIPPHHIIHWYQLEBECANESQAVSVTNPY 480  
DB 421 PQIGEKSLISPVDSYQYGTTLCTVYAIIPPHHIIHWYQLEBECANESQAVSVTNPY 480  
QY 481 PCSEWRSVDFQGNKLEVNKNOPALIEGKNKTVSTLVIQAAVNSALYKCAVNVKGRGE 540  
DB 481 PCSEWRSVDFQGNKLEVNKNOPALIEGKNKTVSTLVIQAAVNSALYKCAVNVKGRGE 540  
QY 541 RVISFHVTRGPETITLQPMQPTQESVSLMCTADRSTFENLTWYKLGPPQLPIHVGLPT 600  
DB 541 RVISFHVTRGPETITLQPMQPTQESVSLMCTADRSTFENLTWYKLGPPQLPIHVGLPT 600  
QY 601 PVCKNLDLTKLNATMFSNSTNDILIMELKNASLODQGDYVCLAQDRKTKKHCVCVRQLT 660  
DB 601 PVCKNLDLTKLNATMFSNSTNDILIMELKNASLODQGDYVCLAQDRKTKKHCVCVRQLT 660  
QY 661 VLVRVAPTIIGNLENOTTISGESIEVSCASGNPPPOIMFKNDETIVESGIVLKDQNR 720  
DB 661 VLVRVAPTIIGNLENOTTISGESIEVSCASGNPPPOIMFKNDETIVESGIVLKDQNR 720  
QY 721 NLTIIRVRKEDEGLYTQACSVLGCACVFAFFIIEGAQEKTNLSIILVGTAVIAMFFWL 780  
DB 721 NLTIIRVRKEDEGLYTQACSVLGCACVFAFFIIEGAQEKTNLSIILVGTAVIAMFFWL 780  
QY 781 LLVILRTVVRANGGELKTYLSIVMDPDELPLDHCERLPYDASKWFFPRDRRLKLGKPL 840  
DB 781 LLVILRTVVRANGGELKTYLSIVMDPDELPLDHCERLPYDASKWFFPRDRRLKLGKPL 840  
QY 841 GRGAFQGVIEADAFIDKATCTRTVAVKMLKEGATHSEHRALMSELKILIHGHLLNVN 900  
DB 841 GRGAFQGVIEADAFIDKATCTRTVAVKMLKEGATHSEHRALMSELKILIHGHLLNVN 900  
QY 901 LLGACTKPGGLMWIVVEFCFKNLSTYLSRKNEFVYKTKGARFRQGDYVGAIPVDLK 960  
DB 901 LLGACTKPGGLMWIVVEFCFKNLSTYLSRKNEFVYKTKGARFRQGDYVGAIPVDLK 960  
QY 961 RRLDSITSSOSSASSGFVEEKSLSDEVEEAPEDLYKDELTLHLICYSFOVAKGMEFLA 1020  
DB 961 RRLDSITSSOSSASSGFVEEKSLSDEVEEAPEDLYKDELTLHLICYSFOVAKGMEFLA 1020  
QY 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPKWMAPETIFDR 1080  
DB 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPKWMAPETIFDR 1080  
QY 1081 VYTIQSDVWFGVLLWEIFSLGASPYGVKIDEEFCRRLKEGTRMRAPDYTTPEMYQTML 1140  
DB 1081 VYTIQSDVWFGVLLWEIFSLGASPYGVKIDEEFCRRLKEGTRMRAPDYTTPEMYQTML 1140

1141 DCWHGSPQRTFSELVEHLGNLLQANAQQDGKDYIVLPISETLSMEEDSGLSLPTSPVS 1200  
1141 DCWHGSPQRTFSELVEHLGNLLQANAQQDGKDYIVLPISETLSMEEDSGLSLPTSPVS 1200  
1201 CMEEEVCDPKFHYDNTAGISQYLNQSKRKRSPVSVKTFEDIPLEPEVKVIPPDDNQDTS 1260  
1201 CMEEEVCDPKFHYDNTAGISQYLNQSKRKRSPVSVKTFEDIPLEPEVKVIPPDDNQDTS 1260  
1261 GMVLASEELKLTEDRTKLSFGGMPVSKRESVASEGSGNQTSQYSGYHSDDDTDTTVYS 1320  
1261 GMVLASEELKLTEDRTKLSFGGMPVSKRESVASEGSGNQTSQYSGYHSDDDTDTTVYS 1320  
1321 SEAEALLKLEIGVQTGSTAQILQPDGTTLSPPV 1356  
1321 SEAEALLKLEIGVQTGSTAQILQPDGTTLSPPV 1356

RESULT 6  
US-09-949-016-9853  
; Sequence 9853, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 9853  
; LENGTH: 1456  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-9853

Query Match 99.9%; Score 7092; DB 2; Length 1456;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 MESKVLAVLWLCVETRAASVGLSPVSLDPLRLSIQKQILTIKANTLQITCRGQDLD 60  
101 MQSKVLLAVLWLCVETRAASVGLSPVSLDPLRLSIQKQILTIKANTLQITCRGQDLD 160

61 WLWPNQSGSEQRVEVTECSDGLFCKTLTIPKVI GNDTGAYKCFYRETDLASVIVYVQD 120  
161 WLWPNQSGSEQRVEVTECSDGLFCKTLTIPKVI GNDTGAYKCFYRETDLASVIVYVQD 220

121 YRSPFTASVSDQGVVYITENKNTVIVPCLSISNLNLSLCARYPEKRPVPGNRI SWD 180  
221 YRSPFTASVSDQGVVYITENKNTVIVPCLSISNLNLSLCARYPEKRPVPGNRI SWD 280

181 SKKGFTIPSYMIYAGVVFCEAKINDESYOSIMYIVVVVGYRDIYVVLSPSHGIELSVGE 240  
281 SKKGFTIPSYMIYAGVVFCEAKINDESYOSIMYIVVVVGYRDIYVVLSPSHGIELSVGE 340

241 KVLNCTARTELNVGIDFNWEPSSKHQKLVNRDLKTQSGSEMKEFLSTLTIDGVTRS 300  
341 KVLNCTARTELNVGIDFNWEPSSKHQKLVNRDLKTQSGSEMKEFLSTLTIDGVTRS 400

301 DQGLYTCAASSGLMTKGNSTFVRVHKPFFVAFSGMESLVEATVGERVRIPAKYLGYPPP 360  
401 DQGLYTCAASSGLMTKGNSTFVRVHKPFFVAFSGMESLVEATVGERVRIPAKYLGYPPP 460

361 EIKWYKNGIPLESNHTIKAGHVLITMEVSRDGTGNVTILTNPISEKQSHVSVLVVYVP 420  
461 EIKWYKNGIPLESNHTIKAGHVLITMEVSRDGTGNVTILTNPISEKQSHVSVLVVYVP 520

421 POIGEXSLISPVDYSQVGTQTTLCTVYAI PPPHHIHWYQLEBEECANEPSQAVSVTNPY 480  
521 POIGEXSLISPVDYSQVGTQTTLCTVYAI PPPHHIHWYQLEBEECANEPSQAVSVTNPY 580  
481 PCSEWRSVEDFOGNGKIEVNKNQFALIEGKNTVSTLVIOAANVSALYKCEAVNKVGRGE 540  
581 PCSEWRSVEDFOGNGKIEVNKNQFALIEGKNTVSTLVIOAANVSALYKCEAVNKVGRGE 640  
541 RVLSFHVTRGPEITLQPDMPQTEQESVSLWCTADRSTFENLTWYKLGPOPLPHVGE LPT 600  
641 RVLSFHVTRGPEITLQPDMPQTEQESVSLWCTADRSTFENLTWYKLGPOPLPHVGE LPT 700  
601 PUCNKLDTLWKLNATWFSNSTDILIMELKNASLOQGDYVCLAQDKTKKRHCVVRLT 660  
701 PUCNKLDTLWKLNATWFSNSTDILIMELKNASLOQGDYVCLAQDKTKKRHCVVRLT 760  
661 VLERVAPTITGNLENQTTSGESIEVSTASGNPPPOIMWFKDNETLVEDSGIVLKDGNR 720  
761 VLERVAPTITGNLENQTTSGESIEVSTASGNPPPOIMWFKDNETLVEDSGIVLKDGNR 820  
721 NLTIRVRKEDGLYTCQACSVLGCACVAFPIIEGAQEKTNLEIILVGTAVIAMFFWL 780  
821 NLTIRVRKEDGLYTCQACSVLGCACVAFPIIEGAQEKTNLEIILVGTAVIAMFFWL 880  
781 LLVIIILRTVKRANGGELKTGYLSIVMDPDELPLDEHCERLPYDASKWEPPDRDLKLGKPL 840  
881 LLVIIILRTVKRANGGELKTGYLSIVMDPDELPLDEHCERLPYDASKWEPPDRDLKLGKPL 940  
841 GRGAFQVIEADAFGIDKTATCRTVAVKMLKEGATHSEHRLMSELKILIHGHILNVVN 900  
941 GRGAFQVIEADAFGIDKTATCRTVAVKMLKEGATHSEHRLMSELKILIHGHILNVVN 1000  
901 LLGACTKPGGPLMVIIEFCKFGLNLTSLRSKRNEFVPIYTKGARFQGDYVGAIPVDLK 960  
1001 LLGACTKPGGPLMVIIEFCKFGLNLTSLRSKRNEFVPIYTKGARFQGDYVGAIPVDLK 1060  
961 RRLDSITSSQSSASSGFVEEKSLSVDEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
1061 RRLDSITSSQSSASSGFVEEKSLSVDEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1120  
1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDDPYVRKGDARLPKKNWAPETIFDR 1080  
1121 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDDPYVRKGDARLPKKNWAPETIFDR 1180  
1081 VYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRLKEGTRMADPYTTPEMYQTM L 1140  
1181 VYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRLKEGTRMADPYTTPEMYQTM L 1240  
1141 DCWHGSPQRTFSELVEHLGNLLQANAQQDGKDYIVLPISETLSMEEDSGLSLPTSPVS 1200  
1241 DCWHGSPQRTFSELVEHLGNLLQANAQQDGKDYIVLPISETLSMEEDSGLSLPTSPVS 1300  
1201 CMEEEVCDPKFHYDNTAGISQYLNQSKRKRSPVSVKTFEDIPLEPEVKVIPPDDNQDTS 1260  
1301 CMEEEVCDPKFHYDNTAGISQYLNQSKRKRSPVSVKTFEDIPLEPEVKVIPPDDNQDTS 1360  
1261 GMVLASEELKLTEDRTKLSFGGMPVSKRESVASEGSGNQTSQYSGYHSDDDTDTTVYS 1320  
1361 GMVLASEELKLTEDRTKLSFGGMPVSKRESVASEGSGNQTSQYSGYHSDDDTDTTVYS 1420  
1321 SEAEALLKLEIGVQTGSTAQILQPDGTTLSPPV 1356  
1421 SEAEALLKLEIGVQTGSTAQILQPDGTTLSPPV 1456

RESULT 7  
US-08-810-116-8  
; Sequence 8, Application US/08810116  
; Patent No. 5768660  
; GENERAL INFORMATION:  
; APPLICANT: Terman, Bruce I.  
; APPLICANT: Carrión, Miguel E.



;; TITLE OF INVENTION: Identification of a No. 57686060el Human Growth  
;; TITLE OF INVENTION: Factor Receptor  
;; NUMBER OF SEQUENCES: 12  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: American Cyanamid Company  
;; STREET: One Cyanamid Plaza  
;; CITY: Wayne  
;; STATE: New Jersey  
;; COUNTRY: U.S.A.  
;; ZIP: 07470  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/810,116  
;; FILING DATE: 25-FEB-1997  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/930,548  
;; FILING DATE: 23-NOV-1992  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Gordon, Alan M.  
;; REGISTRATION NUMBER: 30,637  
;; REFERENCE/DOCKET NUMBER: 31,298-01  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 201-831-3244  
;; TELEFAX: 201-831-3305  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1356 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-810-116-8

Query Match 99.6%; Score 7070; DB 1; Length 1356;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1351; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy	1	MESKVLAVLWLCVETRAASVGLPSVSLDPLRLSIQKIDLTITKANTTLOITCRGQDLD	60
Db	1	MESKVLAVLWLCVETRAASVGLPSVSLDPLRLSIQKIDLTITKANTTLOITCRGQDLD	60
Qy	61	WLWPNNSGSEQRVEVEVTECSDGLFCKTLTIPKVIQNDTGAYKCFYRETDLASVIYVYVD	120
Db	61	WLWPNNSGSEQRVEVEVTECSDGLFCKTLTIPKVIQNDTGAYKCFYRETDLASVIYVYVD	120
Qy	121	YRSPFIASVSDQHGQVYITENKNTVVI PCLGSIISNLNLSLCARYPEKRFVPGNRI SWD	180
Db	121	YRSPFIASVSDQHGQVYITENKNTVVI PCLGSIISNLNLSLCARYPEKRFVPGNRI SWD	180
Qy	181	SKKGFTTIPSYMIYAGWVFCEAKINDESYOSIMYIVVWVGRIYDVVLSHSHGIELSVGE	240
Db	181	SKKGFTTIPSYMIYAGWVFCEAKINDESYOSIMYIVVWVGRIYDVVLSHSHGIELSVGE	240
Qy	241	KLVLNCTARTELNIGIDFNWEPSSKHOKKLVRDLKTQSGSBMKKFLSTLTIDGVTRS	300
Db	241	KLVLNCTARTELNIGIDFNWEPSSKHOKKLVRDLKTQSGSBMKKFLSTLTIDGVTRS	300
Qy	301	DQGLYTCAASSGLMTKKNSTFVRVHEKPFVAFGSGMESLVEATVGERVIRIPAKYLGYPYP	360
Db	301	DQGLYTCAASSGLMTKKNSTFVRVHEKPFVAFGSGMESLVEATVGERVIRIPAKYLGYPYP	360
Qy	361	EIKYKNGIPLESNHTIKAGHLTIMEVSDRTGNTVILTNPISEKQSHVSLVYVYP	420
Db	361	EIKYKNGIPLESNHTIKAGHLTIMEVSDRTGNTVILTNPISEKQSHVSLVYVYP	420
Qy	421	POIGKSLISPDVSQYQGTTLCTVYALPPPHHHIHWYQLEECANEPQAVSVTNPY	480
Db	421	POIGKSLISPDVSQYQGTTLCTVYALPPPHHHIHWYQLEECANEPQAVSVTNPY	480

## RESULT 8

US-07-930-548A-8  
; Sequence 8, Application US/07930548A  
; Patent No. 5861301  
; GENERAL INFORMATION:  
; APPLICANT: Terman, Bruce I.  
; APPLICANT: Carrion, Miguel E.  
; TITLE OF INVENTION: Identification of a No. 5861301el Human Growth  
; TITLE OF INVENTION: Factor Receptor  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: American Cyanamid Company

STREET: One Cyanamid Plaza  
CITY: Wayne  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07470  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/930,548A  
FILING DATE: 23-NOV-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Gordon, Alan M.  
REGISTRATION NUMBER: 30,637  
REFERENCE/DOCKET NUMBER: 31,298-01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-831-3244  
TELEFAX: 201-831-3305  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1356 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-930-548A-8

Query Match 99.6%; Score 7070; DB 1; Length 1356;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1351; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy	1	MESKVLAVLWLCVETRAASVGLPSVSLDPLRLSIQKIDILTIKANTLIQITCRGQDLD	60
Db	1	MESKVLAVLWLCVETRAASVGLPSVSLDPLRLSIQKIDILTIKANTLIQITCRGQDLD	60
Qy	61	WLWPNNSGSEQRVEVEVTECDGLFCKTLTIPIKVGNDTGAYKCFYRETDLASVIVYVQD	120
Db	61	WLWPNNSGSEQRVEVEVTECDGLFCKTLTIPIKVGNDTGAYKCFYRETDLASVIVYVQD	120
Qy	121	YRSPFIASVDQHGVIYITENKNTVVIIPCLGSIINLNSLCARYPEKRFVDPGNRISWD	180
Db	121	YRSPFIASVDQHGVIYITENKNTVVIIPCLGSIINLNSLCARYPEKRFVDPGNRISWD	180
Qy	181	SKGFTIPSYMISYAGWVFCEAKINDESYSQIMYIVVVVGYRIYDVVLSPSHGIELSVGE	240
Db	181	SKGFTIPSYMISYAGWVFCEAKINDESYSQIMYIVVVVGYRIYDVVLSPSHGIELSVGE	240
Qy	241	KLVLNCTARTLNVGIDFNWEPSSKHOKHKLNRDLKTQSGSEMCKFLSTLTIIDGVTRS	300
Db	241	KLVLNCTARTLNVGIDFNWEPSSKHOKHKLNRDLKTQSGSEMCKFLSTLTIIDGVTRS	300
Qy	301	DOGLYTCAASGLMTKKNSTFVRVHEKPPFVAFSGMESLVEATVGERVIRIPAKYLGYPVP	360
Db	301	DOGLYTCAASGLMTKKNSTFVRVHEKPPFVAFSGMESLVEATVGERVIRIPAKYLGYPVP	360
Qy	361	EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTGNTVILTNPISEKQSHVVSIVYVVP	420
Db	361	EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTGNTVILTNPISEKQSHVVSIVYVVP	420
Qy	421	POIGEKSLISPDVSQYGTQTTLCTVYAIIPPHHIIHWYQLEECANEPQAVSVTNPY	480
Db	421	POIGEKSLISPDVSQYGTQTTLCTVYAIIPPHHIIHWYQLEECANEPQAVSVTNPY	480
Qy	481	PCBEWRSVEDFQGNKIEVKNQFALIEGKNKTVSTLVIQAANVSALYKCEAVNKVGRGE	540
Db	481	PCBEWRSVEDFQGNKIEVKNQFALIEGKNKTVSTLVIQAANVSALYKCEAVNKVGRGE	540
Qy	541	RVISFHVTRGPETLQPDMPQTOESVSLNCTADRTFFENLTWYKLGQPLPHVGELEPT	600
Db	541	RVISFHVTRGPETLQPDMPQTOESVSLNCTADRTFFENLTWYKLGQPLPHVGELEPT	600

Qy	601	PVCKNLDTLWKLNATMFSNSTNDILIMELKNASLQDQGDYVCLAQDRKTKKRHCVRQLT	660
Db	601	PVCKNLDTLWKLNATMFSNSTNDILIMELKNASLQDQGDYVCLAQDRKTKKRHCVRQLT	660
Qy	661	VLERVAPTTIGNLENQTTISGESIEVSCASGNPPPOIMWFKONETLVEDSGIVLKDGNR	720
Db	661	VLERVAPTTIGNLENQTTISGESIEVSCASGNPPPOIMWFKONETLVEDSGIVLKDGNR	720
Qy	721	NLTIRVRKEDGLYTCQACSVLGCACVAFVFFIEGAQKTNLEIILVCTAVIAMFFWL	780
Db	721	NLTIRVRKEDGLYTCQACSVLGCACVAFVFFIEGAQKTNLEIILVCTAVIAMFFWL	780
Qy	781	LLVLIILTVKRVANGELKTGYLSIVMDPDELPLDEHCERLPYDASKWEFPRDLNLGKPL	840
Db	781	LLVLIILTVKRVANGELKTGYLSIVMDPDELPLDEHCERLPYDASKWEFPRDLNLGKPL	840
Qy	841	GRGAFQGVTEADAFGIDKTATCTVAVKMLKEGATHSEHRALMSELKILIHIGHLNVVN	900
Db	841	GRGAFQGVTEADAFGIDKTATCTVAVKMLKEGATHSEHRALMSELKILIHIGHLNVVN	900
Qy	901	LLGACTKPGGPLMVIIVEFCCKGNLSTYLSKRNFPYKTKGARFRQGDYVCAIPVDLK	960
Db	901	LLGACTKPGGPLMVIIVEFCCKGNLSTYLSKRNFPYKTKGARFRQGDYVCAIPVDLK	960
Qy	961	RRLDSITSSQSSASSGFEVBEKSLSDVEEBEAPEDLYKDLTLEHLICYSFQVAKGMEFLA	1020
Db	961	RRLDSITSSQSSASSGFEVBEKSLSDVEEBEAPEDLYKDLTLEHLICYSFQVAKGMEFLA	1020
Qy	1021	SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR	1080
Db	1021	SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR	1080
Qy	1081	VYTIQSDVMSFGVLLWEIIFSLGASPYGVKIDEEFCRLKEGTRMRAPDYTTPEMYQTM	1140
Db	1081	VYTIQSDVMSFGVLLWEIIFSLGASPYGVKIDEEFCRLKEGTRMRAPDYTTPEMYQTM	1140
Qy	1141	DCWHGSPSORPFSSELVEHLGNLLOAQQDGYIVLPISETLSMEEDSGLSLTPSVS	1200
Db	1141	DCWHGSPSORPFSSELVEHLGNLLOAQQDGYIVLPISETLSMEEDSGLSLTPSVS	1200
Qy	1201	CMEESEVCPKPHYDNTAGISQYLSQNSKRKSRPVSVKTFEDIPLEPEVKVIPPDDNQDTS	1260
Db	1201	CMEESEVCPKPHYDNTAGISQYLSQNSKRKSRPVSVKTFEDIPLEPEVKVIPPDDNQDTS	1260
Qy	1261	GMVLASEELKTLEDRTKLSPSGMVPKSRSEVASEGNSQTSYQSGYHSDDTDTTVYS	1320
Db	1261	GMVLASEELKTLEDRTKLSPSGMVPKSRSEVASEGNSQTSYQSGYHSDDTDTTVYS	1320
Qy	1321	SEAEALLKLEIGVQTGSTAQILQPDGTTLSPPV 1356	
Db	1321	SEAEALLKLEIGVQTGSTAQILQPDGTTLSPPV 1356	

RESULT 9  
US-08-443-861-2  
; Sequence 2, Application US/08443861  
; Patent No. 5851999  
; GENERAL INFORMATION:  
; APPLICANT: Ullrich, Axel  
; APPLICANT: Risau, Werner  
; APPLICANT: Millauer, Birgit  
; APPLICANT: Gazit, Aviv  
; APPLICANT: Levitzki, Alex  
; TITLE OF INVENTION: Flk-1 Is A Receptor For Vascular  
; TITLE OF INVENTION: Endothelial Growth Factor  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711

```
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/443,861
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/193,829
; FILING DATE: 09-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-060
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212)869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1367 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-443-861-2

Query Match 86.1%; Score 6106.5; DB 1; Length 1367;
Best Local Similarity 85.8%; Pred. No. 0;
Matches 1161; Conservative 71; Mismatches 106; Indels 15; Gaps 3;

QY 1 MESKVLAVLWLCVETRAASVGLPSVSLDPLRLSIQKDLITIKANTLTQITCRGQDLD 60
DB 1 MESKALLAVLWFCVETRAASVGLTGDPLPPKLSLTKDILITILANTLTQITCRGQDLD 60
QY 61 WLPNNQSGSEORVEVTEC--SDGLFCKTLTIPKVIQNDTGAYKCFYRETDLASVIYVY 118
DB 61 WLPNQAORDEERVLVTECGGSDIFCKTLTIPRVNGNDGAYKCSYRDVDIASVYVY 120
QY 119 QDVRSPFIASVSDQGVVYITENKNTWIPCLGISINLVSLCARYPEKRFVPDGNRIS 178
DB 121 RDRSPFIASVSDQGVVYITENKNTWIPCRGISINLVSLCARYPEKRFVPDGNRIS 180
QY 179 WDSKKGTFTPSYMSIYAGMVFCEAKINDESYQSIMYIVVWVYRIYDVLSPHIGIELSV 238
DB 181 WDSIEGFTLPSYMSIYAGMVFCEAKINDESYQSIMYIVVWVYRIYDVLSPHIEIELSA 240
QY 239 GEKLVNCTARTELNVGIDFNWEYPSKHOHKKLVNRDLTKQSGSEMKFLSLTIDGVT 298
DB 241 GEKLVNCTARTELNVGLDFTWHSPPSKSHKKITVNRDVKPPFGTVAKMFLSLTIBSVT 300
QY 299 RSDGLYTCASSGLMTKKSTFVRVHEKPPVAFSGMWSIATVGERVIRIPAKYLGY 358
DB 301 KSDGGEYTCVASSGRMKIKRNTFVRVHTKPIAFSGMKSLVEATVSGVRIPVKYLSY 360
QY 359 PPEIKWKYKNGPIESNHTIKAGHVLITMEVSESDTGNVTLTNPISKQKSHVSVLVY 418
DB 361 APDIKWNRGPIESNTYMTVIGDELITMEVTERDAGNYTLTNPISWEKQSHVSVLVN 420
QY 419 VPPQIGKSLISPVDSYQYGTQTLCTCTVVAIPPHHHIHWYQLEECANEPQAVSVTN 478
DB 421 VPPQIGKALISPMDSYQYGTQTLCTCTVVAIPPHHHIHWYQLEECASYRPGQ---TS 476
QY 479 PYPCEWRVSDFOGKNIEVKNQFALIEGKNKTVSLVIAQANVSALYKCEAVNKGVR 538
DB 477 PYACKERHVEDFOGKNIEVKNQYALIEGKNKTVSLVIAQANVSALYKCEAINKAGR 536
QY 539 GERVISFHVTRGPBITLQPMQPTQESVSLWCTADRTSTENLTWYKGPQPLPIHVGE 598
DB 537 GERVISFHVINGPBITVQPAQPTQESVSLWCTADRTSTENLTWYKLGQATSVMEGS 596
QY 599 PTPVCKNLDLWKLNTMFSNSTNDILIMELKNASLQDQGDYVCLAQDRKTKKRHCYVRQ 658

597 LTPVCKNLDLWKLNTMFSNSTNDILIVAFQNASLQDQGDYVCSAQDKTKKRHCYVRQ 656
659 LTVLERVAPTITGNLENQNTTSIGESIEVSCASGNPPQIMWPKDNSTLVEDSGIVLKG 718
657 LIILERMAPMITGNLENQNTTIGETIEVTPASGNPTPHITWPKDNSTLVEDSGIVLRDG 716
719 NRNLTIIRVRKEDEGLYTCOACSVLGCANKEAEPFIEGAQEKTNLEIILVGTAVIAMFF 778
717 NRNLTIIRVRKEDEGLYTCOACSVLGCANKEAEPFIEGAQEKTNLEIILVGTAVIAMFF 776
779 WLLVILIRTVKXANGELKTYLSIVMDPDELPLDEHCERLPYDASKWFFPRDLKLGK 838
777 WLLVILIRTVKXANGELKTYLSIVMDPDELPLDEHCERLPYDASKWFFPRDLKLGK 836
839 PLGRGAFQGVIEADAFIDKATCTRTVAVOMLKEGATHSEHRALMSELKILIHGHHLNV 898
837 PLGRGAFQGVIEADAFIDKATCTRTVAVOMLKEGATHSEHRALMSELKILIHGHHLNV 896
899 VNLIGACTKPGGPMVIVVECKEKNLSTYLSRGRNEFVPYKTKGAREGQKDYVGAIPVD 958
897 VNLIGACTKPGGPMVIVVECKEKNLSTYLSRGRNEFVPYKTKGAREGQKDYVGAIPVD 956
959 LKRLDSITSSQSSASSGFFVEEKSLSVDEEERAPEDLYKDFLTLEHLICYSFQVAKGMEF 1018
957 LKRLDSITSSQSSASSGFFVEEKSLSVDEEERAPEDLYKDFLTLEHLICYSFQVAKGMEF 1016
1019 LASRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIF 1078
1017 LASRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIF 1076
1079 DRYVTYTSQSDVMSGVLLWEIFSLGASPYPGVKIDEEFCRLKKEGTRMRAPDYTPPEMYQT 1138
1077 DRYVTYTSQSDVMSGVLLWEIFSLGASPYPGVKIDEEFCRLKKEGTRMRAPDYTPPEMYQT 1136
1139 MLDWHEGEPQSRTFSELVHNLQANAAQDGKDVILPISETLSMEBDSGLSLPTSP 1198
1137 MLDWHEGEPQSRTFSELVHNLQANAAQDGKDVILPISETLSMEBDSGLSLPTSP 1196
1199 VSCWEEVEEVCDFPHYDNTAGISQYLSQNSKRKSRPVSVKTFEDIPLEEPVKVIPPDDNQ 1258
1197 VSCWEEVEEVCDFPHYDNTAGISQYLSQNSKRKSRPVSVKTFEDIPLEEPVKVIPPDDNQ 1256
1259 DSGWVLASELKTLEDRTKLSPSGFGMVPSKSRVASEGNSQTSQSGYSHSDDTTTV 1318
1257 DSGWVLASELKTLEDRTKLSPSGFGMVPSKSRVASEGNSQTSQSGYSHSDDTTTV 1316
1319 YSSEAEELKLIEIGVOTGSTAQILQPDSCGTTL 1351
1317 YSSEAEELKLIEIGVOTGSTAQILQPDSCGTTL 1340

RESULT 10
US-08-193-829B-2
; Sequence 2, Application US/08193829B
; Patent No. 6177401
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; APPLICANT: Rieau, Werner
; APPLICANT: Millauer, Birgit
; APPLICANT: Gazit, Aviv
; APPLICANT: Levitzki, Alex
; TITLE OF INVENTION: Flk-1 Is A Receptor For Vascular
; TITLE OF INVENTION: Endothelial Growth Factor
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
```



CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/766,678  
FILING DATE: 25-Jan-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/193,829  
FILING DATE: 09-FEB-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7683-060  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)790-9090  
TELEFAX: (212)869-9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1367 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-766-678-2

Query Match 86.1%; Score 6106.5; DB 2; Length 1367;  
Best Local Similarity 85.8%; Pred. No. 0;  
Matches 1161; Conservative 71; Mismatches 106; Indels 15; Gaps 3;

QY 1 MESKLLAVALWLCVETRAASVGLPSVSLDPLRLISIQKIDILTIKANTTLQITCRGQDLD 60  
DB 1 MESKALLAVALWFCVETRAASVGLTGDFLPPKLSQKIDILTIKANTTLQITCRGQDLD 60  
QY 61 WLMPNNSGSEQRVEVTEC--SDGLFCKTITIPKIVGNDTGAYKCFVRETDLASIVYVY 118  
DB 61 WLMPNAQDSEERVLVTECGGSDSIFCKTILIPRVGNDTGAYKCSVRDVIDIASVYVY 120  
QY 119 ODVRSPIASVSDOHGVVYITENKNTKTVIPCLGISNLNVSLCARYPEKRFVDPGNRI 178  
DB 121 RDVRSPIASVSDOHGVVYITENKNTKTVIPCRGISNLNVSLCARYPEKRFVDPGNRI 180  
QY 179 WDSKKGFTIPSYMSIYAGMVFCEAKINDESYQSIYIMYVWVYRIYDVVLSPSHGIELSV 238  
DB 181 WDSEIGFTLPSYMSIYAGMVFCEAKINDEYQSIYIMYVWVYRIYDVVLSPPHELSA 240  
QY 239 GEKLVNCTARTELNVGIDFNWEYPSKSHQKLVNRDLTKTQSGSEMCKFLSTLTIDGVT 298  
DB 241 GEKLVNCTARTELNVGLDFTWHPSPSKSHHKIKVNRDVKPFPPTVAKMFLSTLTIESVT 300  
QY 299 RSDGLYTCASSGLMTKKNSTFVRVHEKDFVARGSMESLVEATVGERVRIPAKYLGP 358  
DB 301 KSDQGEYTCVASSGRMIKKNRTFVRVHTKPFPIAFGSGMKSLEATVGSQVRIIPVKYLSY 360  
QY 359 PPEIKWYKNGIPLESNHTIKAGHVLITIMEYSDRTGNVTYVILTNPISEKQSHVWSLVVY 418  
DB 361 APDKWYRNRPESNYTMIVGDELITIMEYTERDAGNYTVILTNPISEKQSHVWSLVVN 420  
QY 419 VPPQIGSKLISPVDSYQYGTQTLCTVVAIPPHHIIHWYQLEECANEPQASVVTN 478  
DB 421 VPPQIGKALISPMDSYQYGTQTLCTVVAIPPHHIIHWYQLEECANEPQASVVTN 476  
QY 479 PYPCERHRSVEDPQGNKIEVKNQFALIEGKNKTVSTLVIQANVSALYKCEAVNKVR 538  
DB 477 PYACKERHRSVEDPQGNKIEVKNQYALIEGKNKTVSTLVIQANVSALYKCEAINKVR 536  
QY 539 GERVISPHVTRGPITLQDMPQTEQSSVSIWCTADSTFENLTWYKLGQPLPIHVGL 598  
DB 537 GERVISPHVTRGPITVQAAQTEQSSVSLCTADRTFENLTWYKLGQATSVEMGES 596  
QY 599 PTPVCKNLDLWKLNTWFSNSTNDILIMELKNASLQDQGYVCLAQDRTKKRHCYVRQ 658  
DB 597 LTPVCKNLDLWKLNTWFSNSTNDILIVAFQNASLQDQGYVCSAQDKTKKRHCYVRQ 656  
QY 659 LTVLERVAPITITGNLENQTTSIGESIEVSCVTASGNPPPPQIMWFKDNETLVEDSGIVLKOG 718

Db 657 LILERMAPMITGNLENQTTIGETIEVTCPSGNPTPHITWPKDNETLVEDSGIVLRDG 716  
QY 719 NRNLTIIRVRKEDGLYTCQACSVLGCAYEAPFIIEGAQKTNLEIILVGTAVIAMFF 778  
Db 717 NRNLTIIRVRKEDGLYTCQACSVLGCARAEITFIIEGAQKTNLEIILVGTAVIAMFF 776  
QY 779 WLIIIVILRTVRKANGELTKYLSIVMDPDELPLDHECERLKYDASKWFFPRDLKLGK 838  
Db 777 WLIIIVILRTVRKANGELTKYLSIVMDPDELPLDHECERLKYDASKWFFPRDLKLGK 836  
QY 839 PLGRGAFQVIEADAFGIDKATCTRTVAVQMLKEGATHSEHRALMSLKLIIHGHILNV 898  
Db 837 PLGRGAFQVIEADAFGIDKATCTRTVAVQMLKEGATHSEHRALMSLKLIIHGHILNV 896  
QY 899 VNLGACTKPGGPMVIVBFCKFGNLSYLRSKRNFVVPYKTKGARPRQKQDYVGAIPVD 958  
Db 897 VNLGACTKPGGPMVIVBFCKFGNLSYLRSGRNFVVPYKTKGARPRQKQDYVGAIPVD 956  
QY 959 LKERLDSITSSQSSASSGFFVEEKSLSDVDEEEREPEDLYKDFLTLEHLICYSFQVAKGMEF 1018  
Db 957 LKERLDSITSSQSSASSGFFVEEKSLSDVDEEEREPEDLYKDFLTLEHLICYSFQVAKGMEF 1016  
QY 1019 LASRKCITHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIF 1078  
Db 1017 LASRKCITHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIF 1076  
QY 1079 DRVYTTQSDVMSFGVLLWEIFSLGASPGVKIDEEFCRLKEGTRMRAPDYTTPEMYQT 1138  
Db 1077 DRVYTTQSDVMSFGVLLWEIFSLGASPGVKIDEEFCRLKEGTRMRAPDYTTPEMYQT 1136  
QY 1139 MLDWCHGEPSORTFSELVEHLGNLQANAQQQKDYIVLPISSETLSMEEDSGLSLTSP 1198  
Db 1137 MLDWCHGEPSORTFSELVEHLGNLQANAQQQKDYIVLPISSETLSMEEDSGLSLTSP 1196  
QY 1199 VSCWEEVEEVCDPKFDYDNTAGISYQLNKRKRSPVSVKTFEDIPLPEEPVKVIPPDDNQ 1258  
Db 1197 VSCWEEVEEVCDPKFDYDNTAGISYQLNKRKRSPVSVKTFEDIPLPEEPVKVIPPDDNQ 1256  
QY 1259 DSGMVLASBELKTLTDRTKLSPSFGMVPKSRRESVASEGNSQTSYGQSHSDDTDTTV 1318  
Db 1257 DSGMVLASBELKTLTDRTKLSPSFGMVPKSRRESVASEGNSQTSYGQSHSDDTDTTV 1316  
QY 1319 YSSEBAELKLIEIGVQTSQTAILOPDGTTL 1351  
Db 1317 YSSEBAELKLIEIGVQTSQTAILOPDGTTL 1340

RESULT 12  
US-07-813-593-4  
Sequence 4, Application US/07813593  
Patent No. 5185438  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ibor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMCLONE SYSTEMS INCORPORATED  
STREET: 180 VARICK STREET  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/813,593  
FILING DATE: 19920415  
CLASSIFICATION: 435



PRIOR APPLICATION DATA: US UNASSIGNED  
FILING DATE: 12-NOV-1992  
PRIOR APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
PRIOR APPLICATION DATA: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
PRIOR APPLICATION DATA: TW 81102961  
FILING DATE: 15-APR-1992  
PRIOR APPLICATION DATA: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
PRIOR APPLICATION DATA: US 07/813,593  
FILING DATE: 24-DEC-1991  
PRIOR APPLICATION DATA: US 07/793,065  
FILING DATE: 15-NOV-1991  
PRIOR APPLICATION DATA: US 07/728,913  
FILING DATE: 28-JUN-1991  
PRIOR APPLICATION DATA: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1367 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-977-451-6

Query Match 85.98; Score 6097.5; DB 1; Length 1367;  
Best Local Similarity 85.78; Pred. No. 0;  
Matches 1160; Conservative 72; Mismatches 106; Indels 15; Gaps 3;

QY 1 MESKLLAVALWLCVETRAASVGLPSVSLDPLRLSIQKIDILTIKANTTLQITCRGQDLD 60  
DB 1 MESKGLAVALWFCVETRAASVGLPGDFLHPPKLTSTQKIDILTIANTTLQITCRGQDLD 60  
QY 61 WLPNNSGSEQRVEVTEC--SDGLFCKTLTIPKVIQNDTGAYKCFYRETDLASVIYVVV 118  
DB 61 WLPNQAORDSEERVLVTECGGDSIFCKLTIPRVGNDTGAYKCSYRDVDIASTVYVVV 120  
QY 119 QDYSPIASVSDQGVVYITENKNTWVPCIGSINLNVSLCARYPERKFPDGNRIIS 178  
DB 121 RDYSPIASVSDQGVVYITENKNTWVPCIGSINLNVSLCARYPERKFPDGNRIIS 180  
QY 179 WDSKKGFTPSYMLSYAGMVFCEAKINDESYQSMYIVVVGYRIYDVVLSPSHGIELSV 238  
DB 181 WDSKKGFTPSYMLSYAGMVFCEAKINDETYSYQSMYIVVVGYRIYDVVLSPSHGIELSA 240  
QY 239 GEKLVNCTARTLNVDGFNWEYFSSKHQKHLVNRDLKLTQSGSEMKEFLSTLITDGV 298  
DB 241 GEKLVNCTARTLNVDGFNWEYFSSKHQKHLVNRDLKLTQSGSEMKEFLSTLITDGV 300  
QY 299 RSDQGLYTCASSGLMTKNSFTVRVHEKFPVAFSGSMESLVEATVGERVIRIPAKYLY 358  
DB 301 KSDQGEYTCVASSGRMIRKRNRTFVRVHTKPPAFSGSMKSLVEATVGSQVRIPVKYLY 360  
QY 359 PPEIKWYKNGIPLSNHTIKAGHVLTIMEYSEDTGNTVLTNPISKEQSHVSLAVVY 418  
DB 361 APDIKWYRNGRPIESNTYTMVGDLETIMEVTERDAGNTVLTNPISKEQSHVSLAVN 420

QY 419 VPPQIGKSLISPDVSVOYGTQTTLCTVYVAIPPHHIIHMYWQLEBECANEPQAVSVTN 478  
DB 421 VPPQIGKALISPDMSYQYGTMTLCTVYANPLHHIQWYQLEBECASYRPGQ----TS 476  
QY 479 PYPCEBWSRVEDPQGGNKIEVNKNQFALIEGKNKTSTVLVIQAANSALYKCEAVNKVR 538  
DB 477 PYACKERHVEDPQGGNKIEVTKQYALIEGKNKTSTVLVIQAANSALYKCEAINKAGR 536  
QY 539 GERVISFHVTRGPEITLQPMQPTQESVSLMCTADRTSTPENITWYKLGQPIPIHV 598  
DB 537 GERVISFHVIRGPEITVQPAAPTEQESVSLMCTADRTSTPENITWYKLGQPIPIHV 596  
QY 599 PTVCCKNLDTLWKNATMFSNSTNDILIMELKNASLODQGVYVCLADRTKRRHCVV 658  
DB 597 LTPVCCKNLDTLWKNATMFSNSTNDILIVAFQNASLODQGVYVCLADRTKRRHCVV 656  
QY 659 LTVLVRVAPITITGNLENQTTISGESIEVSTASGNPPQIMWPKDNETLVEDSGV 718  
DB 657 LTVLVRVAPITITGNLENQTTISGESIEVSTASGNPPQIMWPKDNETLVEDSGV 716  
QY 719 NRNLTIIRVRKEDGLYTQACSVLGCARVAFIIEGAQKTNLEIILVGTAVIAMFF 778  
DB 717 NRNLTIIRVRKEDGLYTQACSVLGCARVAFIIEGAQKTNLEIILVGTAVIAMFF 776  
QY 779 WLLVILITVRKANGELKTGYLSIVMDDELPLDEHCERLPYDASKWFFPRDLKLG 838  
DB 777 WLLVILITVRKANGELKTGYLSIVMDDELPLDEHCERLPYDASKWFFPRDLKLG 836  
QY 839 PLGRGAFGOVIEADAFIDKATCRTVAVMLKEGATHSEHRAHMLSELKLIHGHILNV 898  
DB 837 PLGRGAFGOVIEADAFIDKATCRTVAVMLKEGATHSEHRAHMLSELKLIHGHILNV 896  
QY 899 VNLGACTKPGGLMVIIEFCKFNLSTYLSKRNFEVFPYKTKGARFQCKQVYGAIPVD 958  
DB 897 VNLGACTKPGGLMVIIEFCKFNLSTYLSKRNFEVFPYKTKGARFQCKQVYGAIPVD 956  
QY 959 LKRLDSITSSQSSASSGFFVEKSLSDVEEEAPEDLYKDFLIEHLICYSFQVAKMEF 1018  
DB 957 LKRLDSITSSQSSASSGFFVEKSLSDVEEEAPEDLYKDFLIEHLICYSFQVAKMEF 1016  
QY 1019 LASRKCITHRDLAARNILLSEKNVVKICDFGLARDIYKDPYVRKGDARLPLKWMAPTIF 1078  
DB 1017 LASRKCITHRDLAARNILLSEKNVVKICDFGLARDIYKDPYVRKGDARLPLKWMAPTIF 1076  
QY 1079 DRYTTIQSDVMSFGLVLEIFSLGASPYGKIDEEFCRLKEGTRMRAPDYTTPEMYQT 1138  
DB 1077 DRYTTIQSDVMSFGLVLEIFSLGASPYGKIDEEFCRLKEGTRMRAPDYTTPEMYQT 1136  
QY 1139 MLDCHWGEPSQRTFSELVEHLGNLQANAQQDGKDYIVLPISSETLSMEEDSGLSLTSP 1198  
DB 1137 MLDCHWGEPSQRTFSELVEHLGNLQANAQQDGKDYIVLPISSETLSMEEDSGLSLTSP 1196  
QY 1199 VSCWEEVEECVDPKPHYDNTAGISYLONSKRKSRPVSVKTFEDIPLPEEPVKVIPP 1258  
DB 1197 VSCWEEVEECVDPKPHYDNTAGISYLONSKRKSRPVSVKTFEDIPLPEEPVKVIPP 1256  
QY 1259 DSGWVLASELKTLEDRTKLSPSFGGMVPSKRESVASEGNSQTSYGQSHSDDTT 1318  
DB 1257 DSGWVLASELKTLEDRTKLSPSFGGMVPSKRESVASEGNSQTSYGQSHSDDTT 1316  
QY 1319 YSSEAEELKLEIGVQTSQTAQLPDGSGTTL 1351  
DB 1317 YSSEAEELKLEIGVQTSQTAQLPDGSGTTL 1340

RESULT 14  
US-07-946-507-4  
Sequence 4, Application US/07946507  
Patent No. 5283354  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL





```

; GENERAL INFORMATION:
;
; APPLICANT: Lemischka, Ihor R.
;
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
;
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
;
; NUMBER OF SEQUENCES: 10
;
; CORRESPONDENCE ADDRESS:

```

	Query Match	85.9%	Score 6097.5	DB 1	Length 1367
	Best Local Similarity	85.7%	Pred. NO. 0		
	Matches 1160	Conservative 72	Mismatches 106	Indels 15	Gaps 3
Qy	1	MESKVLVALVNLVCVETRAASVGLPSVSLDLPELSIQKOLLTIKANTLQITCRGORDLD	60		
Db	1	MESKGLVALVNLVCVETRAASVGLPGDFLPPKLSIQKOLLTIKANTLQITCRGORDLD	60		
Qy	61	WLVPNNQSGSEQREVEYTEC--SDGLEFKTLTIPKVIIGNDTGAYKCFYRETDLASVIVVV	118		

61	WLPNQRDSERVLVTECGGDSIFCKTLTTPRVVGNDTGAYKCSYRVDVDASTVYVV	120
119	QYDRSPFIASVDQHGQVYVITENKNTKTVIPCLGSIISNLNVSICARYPEKRFVPDGNRIS	178
121	RDYRSPFIASVDQHGIVITENKNTKTVIPCRGSIISNLNVSICARYPEKRFVPDGNRIS	180
179	WDSKGFPTIPSYMISYAGVMVFCEAKINDRESYQSIMYIVVVGVYRIIDVVLSPSHGIELSV	238
181	WDSIEGFTLPSYMISYAGVMVFCEAKINDETYSIMYIVVVGVYRIIDVVLSPHPHIELSA	240
239	GEKVLNCTARTELNVGIDFNWEYPPSKHOKHKLNNRDLKTQSGSWMKFLSTLTIDGVT	298
241	GEKVLNCTARTELNVLGDFTHWSPSPSKSHHKIKVNRDVKPPFGTVAKMFSLTLTIESVT	300
299	RSDQGLYTCAASGLMTKKNSTFVRVHEKFPFAFGSMESLVEATYGERVRIIPAKYLGYP	358
301	KSDQGEYTCVASSGRMICKNRITFVRVHTKFPFAFGSMKSLVEATYGSQVRIIPVKVLSY	360
359	PREIKYKNGIPILESNHTIKAGHVLTIMEVSEBDTGNVTIVILTNPISKEQSHVSVLVY	418
361	APDIKWYRNGRTIESNYTMIVGDELTIMEVTERDAGNYTIVILTNPISMEKQSHVSVLVN	420
419	VPOIGEKSILSPVDSYQGTTLCTTCTVYVAPPHPHIIHWQLEBECANEPQAVSVTN	478
421	VPOIGEKALISPMDSYQGTWTLCTCTVYVAPPHPHIIHWQLEBACSVRPQ-----TS	476
479	PYPCEWRSEVDFQGGNKIEVKNQPALIEGKNKTVSTLVIQAANYSALYKCEAVNKVGR	538
477	PYACKERWHEVDFQGGNKIEVTKQYALIEGKNKTVSTLVIQAANYSALYKCEAINKAGR	536
539	GERVISFHVTRGEITLOPDMOPTQESVSVMCTADRSTFENITWYKLGDPQPIPIHVGEL	598
537	GERVISFHVTRGEITVQPAAPTEQESVSLCTADRNTFENITWYKLGSAQTSVHMGES	596
599	PTPVCKNLDTLKLKNATMESNNDLILMELKNASLODQDGVYVCLADRKTKKRHCVVRO	658
597	LTPVCKNLDAUKLNGTMEFSNNDLILVAFQNASLODQDGVYVCSAQDCKTKKRHCVLVKQ	656
659	LTVLERVAPTTICNLNQTSSIGESIEVSTAGSNPPQIMFKNMETLVEDSGIVLKDG	718
657	LIILERMAPWITCNLENOITIGETIEVTCPASGNTPHITWPKDNETLVEDSGIVLRDG	716
719	NNRLTIRRVKEDGELYTQACSVLGCAKVEAFPIIEGAQEKTNLEIILVGTAVIAMFF	778
717	NNRLTIRRVKEDGELYTQACNVLGCAEAETLPIIEGAQEKTNLEIILVGTAVIAMFF	776
779	WLLLVILTVKVRANGELKTGYLSIVMPDDELPLDEHCERLPDYASKWEPFDRDLKJG	838
777	WLLLVILTVKVRANGELKTGYLSIVMPDDELPLDERCERLPDYASKWEPFDRDLKJG	836
839	PLGRGAFQVIEADAFGIDKTATCTVAVKMLKEGATHSEHRALMSELKILIHGHILNV	898
837	PLGRGAFQVIEADAFGIDKTATCTVAVKMLKEGATHSEHRALMSELKILIHGHILNV	896
899	VNLLGACTKPGGPLMVIIEFCKFGNLTLYLSKRNEFPYKTKGARFQSGKVYGAIPVD	958
897	VNLLGACTKPGGPLMVIIEFSGFGNLTLYLRGRNEFPYKSKGARFQSGKVYGBLSVD	956
959	LKRRLDITSOSSASSGFVEEKSLSDVDEEERAPEDLYKDFLTLEHLICYSFOVAKGMEF	1018
957	LKRRLDITSOSSASSGFVEEKSLSDVDEEERASEELYKDFLTLEHLICYSFOVAKGMEF	1016
1019	LASRKCIHRDLAARNLILSEKNVVKICDFGLARDIYKDDPYVRKGDARLPKWMAPETIF	1078
1017	LASRKCIHRDLAARNLILSEKNVVKICDFGLARDIYKDDPYVRKGDARLPKWMAPETIF	1076
1079	DRVYTTIQSDVWSFGVLLWEIFSLGASPYGVKIDEEFCRRLEKGTMRAPDYTTPEMYQT	1138
1077	DRVYTTIQSDVWSFGVLLWEIFSLGASPYGVKIDEEFCRRLEKGTMRAPDYTTPEMYQT	1136
1139	MLDCWHGEPQRSPTFSELVEHLGNLLQANAOQDGKDIYVLPISETISMEBDSGLSLPTSP	1198
1137	MLDCWHEDNQRPSEFSELVEHLGNLLQANAOQDGKDIYVLPMSETISMEBDSGLSLPTSP	1196

Qy	1199	VSCMEEEVCDPKFHYDNTAGISQYLONSKRKSRPVSVKTFEDIPLEEPEVKVIPDDNOT	1258
Db	1197	VSCMEEEVCDPKFHYDNTAGISHYLQNSKRKSRPVSVKTFEDIPLEEPEVKVIPDDNOT	1256
Qy	1259	DSGMVLASEELKTLEDRTKLSFPGMVPSKRESVASEGSGNOTSGYQSGYHSDDTDTTV	1318
Db	1257	DSGMVLASEELKTLEDRTKLSFPGMMPKRESVASEGSGNOTSGYQSGYHSDDTDTTV	1316
Qy	1319	YSSFEAEULLKLEIGVQTGSTAQILOPDSGTTL	1351
Db	1317	YSDAAGLLKMWDAVHA-----DSGTTL	1340

Search completed: March 10, 2006, 18:39:45  
Job time : 39.7622 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2006, 18:59:00 ; Search time 133.671 Seconds  
(without alignments)  
4238.599 Million cell updates/sec

Title: US-10-633-742-6

Perfect score: 7095

Sequence: 1 MESKVLAVLWLCVETRAA.....GSTAQILQPDSTGLSSPPV 1356

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2\_5/ptodata/1/pubaa/US07\_PUBCOMB.pap:\*  
2: /cgn2\_6/ptodata/1/pubaa/US08\_PUBCOMB.pap:\*  
3: /cgn2\_6/ptodata/1/pubaa/US09\_PUBCOMB.pap:\*  
4: /cgn2\_6/ptodata/1/pubaa/US10A\_PUBCOMB.pap:\*  
5: /cgn2\_6/ptodata/1/pubaa/US10B\_PUBCOMB.pap:\*  
6: /cgn2\_6/ptodata/1/pubaa/US11\_PUBCOMB.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7095	100.0	1356	4	US-10-022-939-2
2	7095	100.0	1356	4	US-10-100-405A-2
3	7095	100.0	1356	4	US-10-327-414-6
4	7095	100.0	1356	4	US-10-165-193A-11
5	7092	99.9	1356	4	US-10-090-183-2
6	7092	99.9	1356	4	US-10-394-322A-66
7	7092	99.9	1356	4	US-10-440-464-129
8	7092	99.9	1356	5	US-10-783-528-61
9	7092	99.9	1356	5	US-10-872-198-115
10	7092	99.9	1356	5	US-10-741-600-1469
11	7092	99.9	1356	5	US-10-741-600-1471
12	7092	99.9	1356	5	US-10-926-806-10
13	7092	99.9	1356	5	US-10-824-982-2
14	7092	99.9	1356	6	US-11-021-951-115
15	7091	99.9	1356	3	US-09-969-037-7
16	7091	99.9	1356	5	US-10-763-276-7
17	7060	99.5	1354	4	US-10-262-538-30
18	7060	99.5	1354	4	US-10-669-176-30
19	6806	95.9	1305	5	US-10-741-600-1470
20	6127.5	86.4	1345	4	US-10-090-183-6
21	6106.5	86.1	1367	3	US-09-766-678-2
22	6106.5	86.1	1367	4	US-10-165-193A-10
23	6106.5	86.1	1367	5	US-10-799-782-2
24	6097.5	85.9	1367	3	US-09-919-408-6
25	6097.5	85.9	1367	3	US-09-872-136-6
26	6097.5	85.9	1367	3	US-10-639-603-6
27	6097.5	85.9	1367	6	US-11-030-539-6

28	4117	58.0	789	4	US-10-101-018-15	Sequence 15, Appli
29	4056	57.2	773	4	US-10-364-949-4	Sequence 4, Appli
30	4017	56.6	767	4	US-10-105-901-2	Sequence 2, Appli
31	4012	56.5	764	4	US-10-091-300-85	Sequence 85, Appli
32	4012	56.5	764	5	US-10-482-630-137	Sequence 137, App
33	4012	56.5	764	5	US-10-506-997-85	Sequence 85, Appli
34	3907	55.1	942	4	US-10-449-609-8	Sequence 8, Appli
35	3890	54.8	738	4	US-10-425-668-34	Sequence 34, Appli
36	3491	49.2	664	4	US-10-101-018-13	Sequence 13, Appli
37	3453	48.7	806	3	US-09-766-678-5	Sequence 5, Appli
38	3453	48.7	806	5	US-10-799-782-5	Sequence 5, Appli
39	3437	48.4	805	4	US-10-036-869-34	Sequence 34, Appli
40	3106	43.8	731	4	US-10-364-949-2	Sequence 2, Appli
41	2958	41.7	567	4	US-10-327-414-8	Sequence 8, Appli
42	2804	39.5	1363	3	US-09-375-248-19	Sequence 19, Appli
43	2804	39.5	1363	4	US-10-661-740-19	Sequence 19, Appli
44	2804	39.5	1363	5	US-10-473-127-951	Sequence 951, App
45	2797.5	39.4	1368	4	US-10-105-901-34	Sequence 34, Appli

## ALIGNMENTS

### RESULT 1

US-10-022-939-2  
; Sequence 2, Application US/10022939  
; Publication No. US20030032160A1  
; GENERAL INFORMATION:  
; APPLICANT: Kendall, Richard L.  
; APPLICANT: Thomas, Kenneth A.  
; APPLICANT: Mao, Xianzhi  
; APPLICANT: Tebben, Andrew  
; TITLE OF INVENTION: HUMAN RECEPTOR TYROSINE KINASE, KDR  
; FILE REFERENCE: 19963YDB  
; CURRENT APPLICATION NUMBER: US/10/022,939  
; CURRENT FILING DATE: 2001-12-18  
; PRIOR APPLICATION NUMBER: 09/483,539  
; PRIOR FILING DATE: 2000-01-14  
; PRIOR APPLICATION NUMBER: 09/098,707  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/050,962  
; PRIOR FILING DATE: 1997-06-18  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1356  
; TYPE: PRT  
; ORGANISM: Human  
US-10-022-939-2

Query Match					100.0%; Score 7095; DB 4; Length 1356;
Best Local Similarity					100.0%; Pred. No. 0;
Matches 1356; Conservative					0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	MESKVLAVLWLCVETRAASVGLPSVSLDPLSLSTQKDLTKANTTLQTCRGQDLD	60		
Db	1	MESKVLAVLWLCVETRAASVGLPSVSLDPLSLSTQKDLTKANTTLQTCRGQDLD	60		
Qy	61	WLWPNQSGSEQRVEVTECSGLFKCTLTTPKIVGNDTGAYKCFYRETDLASVIYVVD	120		
Db	61	WLWPNQSGSEQRVEVTECSGLFKCTLTTPKIVGNDTGAYKCFYRETDLASVIYVVD	120		
Qy	121	YRSPFIASVSDQHGVIYITENKNTVIVPCGISINLVSLCARYPEKRVDPGNRISWD	180		
Db	121	YRSPFIASVSDQHGVIYITENKNTVIVPCGISINLVSLCARYPEKRVDPGNRISWD	180		
Qy	181	SKGFTIPSMISYAGKMFCEAKINDESQSIYIVVWGYRIYDVVLSHGHIELSVGE	240		
Db	181	SKGFTIPSMISYAGKMFCEAKINDESQSIYIVVWGYRIYDVVLSHGHIELSVGE	240		
Qy	241	KLVLNCTARTELNVGIDFNWEYPSKHQKLVNRDLKTQSGSEMKKFLSTLTIDGVTRS	300		
Db	241	KLVLNCTARTELNVGIDFNWEYPSKHQKLVNRDLKTQSGSEMKKFLSTLTIDGVTRS	300		

QY 301 DQGLYTCAASGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIIPAKYLGYPVP 360  
Db 301 DQGLYTCAASGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIIPAKYLGYPVP 360  
QY 361 EIKWYKNGIPLSNHTIKAGHVLTIMEVSRDGTGNVTILNPISEKQSHVSLVYVP 420  
Db 361 EIKWYKNGIPLSNHTIKAGHVLTIMEVSRDGTGNVTILNPISEKQSHVSLVYVP 420  
QY 421 PQIGKSLISPVDSYQGTQTTLCTVYAI PPPHHIHWYQLEEECANESQAVSTNYP 480  
Db 421 PQIGKSLISPVDSYQGTQTTLCTVYAI PPPHHIHWYQLEEECANESQAVSTNYP 480  
QY 481 PCEWRSVEDFQGGNKIEVNKNQFALLIEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE 540  
Db 481 PCEWRSVEDFQGGNKIEVNKNQFALLIEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE 540  
QY 541 RVISFHVTRGPETILOPDMQPTQESVSLWCTADRSTFENLTWYKLGQPLPIHVGE LPT 600  
Db 541 RVISFHVTRGPETILOPDMQPTQESVSLWCTADRSTFENLTWYKLGQPLPIHVGE LPT 600  
QY 601 PCKNLDTLKLNATFMSNTDILLMELKNASI QDQGDYVCLAQDRKTKRHCVRQLT 660  
Db 601 PCKNLDTLKLNATFMSNTDILLMELKNASIQDQGDYVCLAQDRKTKRHCVRQLT 660  
QY 661 VLERVAPITIGNLENQTTISGESIEVSTASGNPPQIMWFKNDETILVEDSGIVLKDGNR 720  
Db 661 VLERVAPITIGNLENQTTISGESIEVSTASGNPPQIMWFKNDETILVEDSGIVLKDGNR 720  
QY 721 NLTIIRVRKDEGLYTQACSVLGCACKVEAPFTIEGAQEKTNLEIILVGTAVIAMPFWL 780  
Db 721 NLTIIRVRKDEGLYTQACSVLGCACKVEAPFTIEGAQEKTNLEIILVGTAVIAMPFWL 780  
QY 781 LLVILLTRKANGELKTCVLSVMDPDELPLDEHCERLPYDASKWEPRDRKLKGKPL 840  
Db 781 LLVILLTRKANGELKTCVLSVMDPDELPLDEHCERLPYDASKWEPRDRKLKGKPL 840  
QY 841 GRGAFQVIEADAFGIDKTATCRTVAVKMLKEGATHSEHRLMSELKILIHIGHLNVVN 900  
Db 841 GRGAFQVIEADAFGIDKTATCRTVAVKMLKEGATHSEHRLMSELKILIHIGHLNVVN 900  
QY 901 LLGACTKPGPLMVIVFECFKNLSTVLRKRNFPYKTKGARFRQGDYVGAIPVDLK 960  
Db 901 LLGACTKPGPLMVIVFECFKNLSTVLRKRNFPYKTKGARFRQGDYVGAIPVDLK 960  
QY 961 RRLDSITSSOSSASSGFVEEKSLSDEVEEAPEDLYKDLTLEHLICYSQVAKGMEFLA 1020  
Db 961 RRLDSITSSOSSASSGFVEEKSLSDEVEEAPEDLYKDLTLEHLICYSQVAKGMEFLA 1020  
QY 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWWAPETIFDR 1080  
Db 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWWAPETIFDR 1080  
QY 1081 VYTIQSDVWSFGVLLWEIFSLGASPGVKIDBEFCRRLKEGTRMAADYTTPEMYQTML 1140  
Db 1081 VYTIQSDVWSFGVLLWEIFSLGASPGVKIDBEFCRRLKEGTRMAADYTTPEMYQTML 1140  
QY 1141 DCWHGSPSQPTSELVEHLGNLQANAQDGDYVLPISETLSWEEDSGLSLTPSPVS 1200  
Db 1141 DCWHGSPSQPTSELVEHLGNLQANAQDGDYVLPISETLSWEEDSGLSLTPSPVS 1200  
QY 1201 CMEEEVCDPKFYHDNTAGISQYLQNSKRKSRPVSVKTFEDIPLEBPEVKVPI PDDNQDTS 1260  
Db 1201 CMEEEVCDPKFYHDNTAGISQYLQNSKRKSRPVSVKTFEDIPLEBPEVKVPI PDDNQDTS 1260  
QY 1261 GMVLASEELKTLERDTKLSFGMVPKSRRESVASEGNSQTSYGQSHSDDTDTTVYS 1320  
Db 1261 GMVLASEELKTLERDTKLSFGMVPKSRRESVASEGNSQTSYGQSHSDDTDTTVYS 1320  
QY 1321 SEEAELLKLEIGVQTSQAIILOPDSGTTLSPPV 1356  
Db 1321 SEEAELLKLEIGVQTSQAIILOPDSGTTLSPPV 1356

RESULT 2  
US-10-100-405A-2  
; Sequence 2, Application US/10100405A  
; Publication No. US20030055239A1  
; GENERAL INFORMATION:  
; APPLICANT: Kendall, Richard L.  
; APPLICANT: Thomas, Kenneth A.  
; APPLICANT: Mao, Xianzhi  
; APPLICANT: Tebben, Andrew  
; TITLE OF INVENTION: HUMAN RECEPTOR TYROSINE KINASE, KDR  
; FILE REFERENCE: 19963YDC  
; CURRENT APPLICATION NUMBER: US/10/100,405A  
; PRIOR FILING DATE: 2002-08-13  
; PRIOR APPLICATION NUMBER: 10/022,939  
; PRIOR FILING DATE: 2001-12-18  
; PRIOR APPLICATION NUMBER: 09/483,539  
; PRIOR FILING DATE: 2000-01-14  
; PRIOR APPLICATION NUMBER: 09/098,707  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/050,962  
; PRIOR FILING DATE: 1997-06-18  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1356  
; TYPE: PRT  
; ORGANISM: Human  
US-10-100-405A-2

Query Match 100.0%; Score 7095; DB 4; Length 1356;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MESKULLAVALMLCVETRAASVGLPSVSLDRLSLTQKDLILIKANTTTQITCRGORDLD 60  
Db 1 MESKULLAVALMLCVETRAASVGLPSVSLDRLSLTQKDLILIKANTTTQITCRGORDLD 60  
QY 61 WLWPNQSGSEORVEVTECSGDLFCCKTLTI PKVIGNDTGAYKCFYRETDLASVIVYVQD 120  
Db 61 WLWPNQSGSEORVEVTECSGDLFCCKTLTI PKVIGNDTGAYKCFYRETDLASVIVYVQD 120  
QY 121 YRSPFIASVSDQGVVYITENKNKTVIIPCLGSIINLANSVLCARYPEKRFVDPDGNRISMD 180  
Db 121 YRSPFIASVSDQGVVYITENKNKTVIIPCLGSIINLANSVLCARYPEKRFVDPDGNRISMD 180  
QY 181 SKKGFTIPSMISYAGWFCFAKINDESYQSIYIVVVVGYRIYDVVLSFGHIELSVGE 240  
Db 181 SKKGFTIPSMISYAGWFCFAKINDESYQSIYIVVVVGYRIYDVVLSFGHIELSVGE 240  
QY 241 KLVLNCTARTELNVGIDFNWEPSSKHQHKLVNRDLKTQSGSEMKKFLSTLTIDGVTRS 300  
Db 241 KLVLNCTARTELNVGIDFNWEPSSKHQHKLVNRDLKTQSGSEMKKFLSTLTIDGVTRS 300  
QY 301 DQGLYTCAASGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIIPAKYLGYPVP 360  
Db 301 DQGLYTCAASGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIIPAKYLGYPVP 360  
QY 361 EIKWYKNGIPLSNHTIKAGHVLTIMEVSRDGTGNVTILNPISEKQSHVSLVYVP 420  
Db 361 EIKWYKNGIPLSNHTIKAGHVLTIMEVSRDGTGNVTILNPISEKQSHVSLVYVP 420  
QY 421 PQIGKSLISPVDSYQGTQTTLCTVYAI PPPHHIHWYQLEEECANESQAVSTNYP 480  
Db 421 PQIGKSLISPVDSYQGTQTTLCTVYAI PPPHHIHWYQLEEECANESQAVSTNYP 480  
QY 481 PCEWRSVEDFQGGNKIEVNKNQFALLIEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE 540  
Db 481 PCEWRSVEDFQGGNKIEVNKNQFALLIEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE 540  
QY 541 RVISFHVTRGPETILOPDMQPTQESVSLWCTADRSTFENLTWYKLGQPLPIHVGE LPT 600  
Db 541 RVISFHVTRGPETILOPDMQPTQESVSLWCTADRSTFENLTWYKLGQPLPIHVGE LPT 600

Qy 601 PVCKNLDLTKLNATMFSTNDILIMELKNASLODQGDYVCLADRKTKRHCVRQLT 660  
Db 601 PVCKNLDLTKLNATMFSTNDILIMELKNASLODQGDYVCLADRKTKRHCVRQLT 660  
Qy 661 VLVRVAPITIGNLENQTTISGESIEVSTAGNPPQIMFPMKNETLVEDSGIVLKDGNR 720  
Db 661 VLVRVAPITIGNLENQTTISGESIEVSTAGNPPQIMFPMKNETLVEDSGIVLKDGNR 720  
Qy 721 NLATIRVRKEDEGLYTQACSVLCAKVEAFFIIEGAQKTNLEIIILVGTAVIAMPFWL 780  
Db 721 NLATIRVRKEDEGLYTQACSVLCAKVEAFFIIEGAQKTNLEIIILVGTAVIAMPFWL 780  
Qy 781 LLVILITVRKANGGELKTGYLSIWMDDPLDEHCEPLPYDASKWEFFPRDLKLGKPL 840  
Db 781 LLVILITVRKANGGELKTGYLSIWMDDPLDEHCEPLPYDASKWEFFPRDLKLGKPL 840  
Qy 841 GRGAFQGVIEADAFGIDKTATCRVAVKMLKEGATHSEHRAHLMSELKILIHGHILNVN 900  
Db 841 GRGAFQGVIEADAFGIDKTATCRVAVKMLKEGATHSEHRAHLMSELKILIHGHILNVN 900  
Qy 901 LLGACTKPGGPLMVI VEFCKFNGNLSTYLSKRNEFVYKTKGARFRQGDYVGAIPVDLK 960  
Db 901 LLGACTKPGGPLMVI VEFCKFNGNLSTYLSKRNEFVYKTKGARFRQGDYVGAIPVDLK 960  
Qy 961 RRLDSITSSOSSASSGFVEEKSLSDBVEEAPEDLYKDFLTLEHLICYSQVAKMEFLA 1020  
Db 961 RRLDSITSSOSSASSGFVEEKSLSDBVEEAPEDLYKDFLTLEHLICYSQVAKMEFLA 1020  
Qy 1021 SRKCIHRDLAARNLLSEKNVVKICDFGLARDIYKDPYVRKGDARLPLKMAPEITFDR 1080  
Db 1021 SRKCIHRDLAARNLLSEKNVVKICDFGLARDIYKDPYVRKGDARLPLKMAPEITFDR 1080  
Qy 1081 VYTIQSDWMSFGVLLWEIFSLGASPYGVKIDIEFCRLKEGTRMRAPDYTPPMYQTM 1140  
Db 1081 VYTIQSDWMSFGVLLWEIFSLGASPYGVKIDIEFCRLKEGTRMRAPDYTPPMYQTM 1140  
Qy 1141 DCWHGSPQRTFSELVEHLGNLLOANAQDGKDYI VLPISSETLSMBEDSGLSPTSPVS 1200  
Db 1141 DCWHGSPQRTFSELVEHLGNLLOANAQDGKDYI VLPISSETLSMBEDSGLSPTSPVS 1200  
Qy 1201 CMEEEVCDPKFHYDNTAGISQYLONSKRKSRPVSVKTFEDIPLREEPVKVI PDDNQTDS 1260  
Db 1201 CMEEEVCDPKFHYDNTAGISQYLONSKRKSRPVSVKTFEDIPLREEPVKVI PDDNQTDS 1260  
Qy 1261 GMVLASBELKTLEDRTKLSPSGMGVPSKRESVASSEGSNOTSGYQSGYHSDDDTTVYS 1320  
Db 1261 GMVLASBELKTLEDRTKLSPSGMGVPSKRESVASSEGSNOTSGYQSGYHSDDDTTVYS 1320  
Qy 1321 SEEAELKLIEIGVGTGSTAQILOPDSGTTLSPPV 1356  
Db 1321 SEEAELKLIEIGVGTGSTAQILOPDSGTTLSPPV 1356

## RESULT 3

US-10-327-414-6  
; Sequence 6, Application US/10327414  
; Publication No. US20030158083A1  
; GENERAL INFORMATION:  
; APPLICANT: The Procter & Gamble Company  
; APPLICANT: Peters, Kevin G  
; TITLE OF INVENTION: A Method of Effecting angiogenesis by Modulating the Function of  
; TITLE OF INVENTION: Endothelial Phosphatase  
; FILE REFERENCE: 8864M  
; CURRENT APPLICATION NUMBER: US/10/327,414  
; CURRENT FILING DATE: 2002-12-20  
; PRIOR APPLICATION NUMBER: US 60/355,125  
; PRIOR FILING DATE: 2002-02-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 6  
; LENGTH: 1356  
; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-327-414-6

Query Match 100.0%; Score 7095; DB 4; Length 1356;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESKVLLAVALWLCVETRAASVGLPSVSLDPLRLSIQDKDILTIKANTTLOITCRGQDLD 60  
Db 1 MESKVLLAVALWLCVETRAASVGLPSVSLDPLRLSIQDKDILTIKANTTLOITCRGQDLD 60  
Qy 61 WLPNNOSGEQREVEVECTSDGLFCKTLTI PKVIGNDTGAYKCFYRETDLASVIVVQD 120  
Db 61 WLPNNOSGEQREVEVECTSDGLFCKTLTI PKVIGNDTGAYKCFYRETDLASVIVVQD 120  
Qy 121 YRSPFIASVSDQHGVIYITENKNTVVI PCIGSISNLSVLCARYPERKFPVDPGNRISWD 180  
Db 121 YRSPFIASVSDQHGVIYITENKNTVVI PCIGSISNLSVLCARYPERKFPVDPGNRISWD 180  
Qy 181 SKKGFTIPSYMI SYAGMVFCEAKINDESYSIMYIVVVGVYRIYDVVLSPSHGIELSVGE 240  
Db 181 SKKGFTIPSYMI SYAGMVFCEAKINDESYSIMYIVVVGVYRIYDVVLSPSHGIELSVGE 240  
Qy 241 KLVLNCTARTELNVDGIDFNWEYPSKHQKHLNRLDKTQSGSEMKKFLSTLTIDGVTRS 300  
Db 241 KLVLNCTARTELNVDGIDFNWEYPSKHQKHLNRLDKTQSGSEMKKFLSTLTIDGVTRS 300  
Qy 301 DOGLYTCAASGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVIRIPAKYLGYP 360  
Db 301 DOGLYTCAASGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVIRIPAKYLGYP 360  
Qy 361 EIKWYKNGIPIESNHTIKAGHVLTIMEVSRDGTGYVTILNTPISKQSHVSVLVVTP 420  
Db 361 EIKWYKNGIPIESNHTIKAGHVLTIMEVSRDGTGYVTILNTPISKQSHVSVLVVTP 420  
Qy 421 PQIGESLIPVDSYQGTITLTCTVYAI PPHHIIHWYQWLBECANESQAVSVNTPY 480  
Db 421 PQIGESLIPVDSYQGTITLTCTVYAI PPHHIIHWYQWLBECANESQAVSVNTPY 480  
Qy 481 PCBWRSVEDFOGKNKIEVNKNQFALIEGKNKTVSTLVIQAANVSALYKCAVNKVG 540  
Db 481 PCBWRSVEDFOGKNKIEVNKNQFALIEGKNKTVSTLVIQAANVSALYKCAVNKVG 540  
Qy 541 RVISFHVTRPETITLOPDMQPTQESVSLMCTADRSFENLTWYKLGQPLPIHVGSLPT 600  
Db 541 RVISFHVTRPETITLOPDMQPTQESVSLMCTADRSFENLTWYKLGQPLPIHVGSLPT 600  
Qy 601 PVCKNLDLTKLNATMFSTNDILIMELKNASLODQGDYVCLADRKTKRHCVRQLT 660  
Db 601 PVCKNLDLTKLNATMFSTNDILIMELKNASLODQGDYVCLADRKTKRHCVRQLT 660  
Qy 661 VLVRVAPITIGNLENQTTISGESIEVSTAGNPPQIMFPMKNETLVEDSGIVLKDGNR 720  
Db 661 VLVRVAPITIGNLENQTTISGESIEVSTAGNPPQIMFPMKNETLVEDSGIVLKDGNR 720  
Qy 721 NLATIRVRKEDEGLYTQACSVLCAKVEAFFIIEGAQKTNLEIIILVGTAVIAMPFWL 780  
Db 721 NLATIRVRKEDEGLYTQACSVLCAKVEAFFIIEGAQKTNLEIIILVGTAVIAMPFWL 780  
Qy 781 LLVILITVRKANGGELKTGYLSIWMDDPLDEHCEPLPYDASKWEFFPRDLKLGKPL 840  
Db 781 LLVILITVRKANGGELKTGYLSIWMDDPLDEHCEPLPYDASKWEFFPRDLKLGKPL 840  
Qy 841 GRGAFQGVIEADAFGIDKTATCRVAVKMLKEGATHSEHRAHLMSELKILIHGHILNVN 900  
Db 841 GRGAFQGVIEADAFGIDKTATCRVAVKMLKEGATHSEHRAHLMSELKILIHGHILNVN 900  
Qy 901 LLGACTKPGGPLMVI VEFCKFNGNLSTYLSKRNEFVYKTKGARFRQGDYVGAIPVDLK 960  
Db 901 LLGACTKPGGPLMVI VEFCKFNGNLSTYLSKRNEFVYKTKGARFRQGDYVGAIPVDLK 960  
Qy 961 RRLDSITSSOSSASSGFVEEKSLSDBVEEAPEDLYKDFLTLEHLICYSQVAKMEFLA 1020  
Db 961 RRLDSITSSOSSASSGFVEEKSLSDBVEEAPEDLYKDFLTLEHLICYSQVAKMEFLA 1020

Db 961 RRLDSITSSQSSASSGFVEEKSLSLSDVEEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
Qy 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDDYVRKGDARPLKWMAPETIFDR 1080  
Db 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDDYVRKGDARPLKWMAPETIFDR 1080  
Qy 1081 VYTIQSDWMSFGVLLWEIFSLGASPYPGVKIDEEFCRLKEGTRMRAPDYTTPEMYQTML 1140  
Db 1081 VYTIQSDWMSFGVLLWEIFSLGASPYPGVKIDEEFCRLKEGTRMRAPDYTTPEMYQTML 1140  
Qy 1141 DCWHGEPSPRTFSELVEHLGNLLQANAQDGKDIYVLPISSETLSMEEDSGLSPTSPTS 1200  
Db 1141 DCWHGEPSPRTFSELVEHLGNLLQANAQDGKDIYVLPISSETLSMEEDSGLSPTSPTS 1200  
Qy 1201 CMEEEVCDPKFHYDNTAGISQYLQNSKRKSRPVSVKTFEDIPLEPEVKVIPPDDNQDTS 1260  
Db 1201 CMEEEVCDPKFHYDNTAGISQYLQNSKRKSRPVSVKTFEDIPLEPEVKVIPPDDNQDTS 1260  
Qy 1261 GMLVASEELKTLEDRTKLSFSGMVPSKRSRESVASEGNSQTSQYQSGYHSDDDTDTTVYS 1320  
Db 1261 GMLVASEELKTLEDRTKLSFSGMVPSKRSRESVASEGNSQTSQYQSGYHSDDDTDTTVYS 1320  
Qy 1321 SEAEALLKLIETGVGTGTAQILQPDGSGTTLSSPPV 1356  
Db 1321 SEAEALLKLIETGVGTGTAQILQPDGSGTTLSSPPV 1356

RESULT 4  
US-10-165-193A-11  
; Sequence 11, Application US/10165193A  
; Publication No. US20030207391A1  
; GENERAL INFORMATION:  
; APPLICANT: HELEN PAPPAS  
; TITLE OF INVENTION: BINDING PROTEIN  
; FILE REFERENCE: 1396-1-00  
; CURRENT FILING DATE: 2003-01-13  
; PRIOR FILING DATE: 2003-01-13  
; PRIOR FILING DATE: 2000-12-07  
; PRIOR FILING DATE: 2000-12-07  
; PRIOR FILING DATE: 1999-12-07  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: SeqWin99, version 1.02  
; SEQ ID NO 11  
; LENGTH: 1356  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-165-193A-11

Query Match 100.0%; Score 7095; DB 4; Length 1356;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESKVLALVALWCVETRAASVGLPSVSLDPLRLSIQKDLTIKANTTLQITCRGQDLD 60  
Db 1 MESKVLALVALWCVETRAASVGLPSVSLDPLRLSIQKDLTIKANTTLQITCRGQDLD 60  
Qy 61 WLWPNNGSGEQRVEVTECSGDLFCCKTLTTPKVI GNDTGAYKCFYRETDLASVIYVYVD 120  
Db 61 WLWPNNGSGEQRVEVTECSGDLFCCKTLTTPKVI GNDTGAYKCFYRETDLASVIYVYVD 120  
Qy 121 YRSPFTASVDQHGVIYITENKNTVYICLGSISNLNLSLCARYPEKRPVDPGNRISWD 180  
Db 121 YRSPFTASVDQHGVIYITENKNTVYICLGSISNLNLSLCARYPEKRPVDPGNRISWD 180  
Qy 181 SKKGFTIPSMYSYAGMVFCEAKINDESYQSIYVVVVGYRIYDVVLSPSHGIELSVGE 240  
Db 181 SKKGFTIPSMYSYAGMVFCEAKINDESYQSIYVVVVGYRIYDVVLSPSHGIELSVGE 240  
Qy 241 KLVNCTARTELNVGIDFNWEPSSKHQKLVNRDLKTQSGSEMKKFLSTLIDGVTRS 300  
Db 241 KLVNCTARTELNVGIDFNWEPSSKHQKLVNRDLKTQSGSEMKKFLSTLIDGVTRS 300

Qy 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGBRVIRIPAKYLGYPVP 360  
Db 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGBRVIRIPAKYLGYPVP 360  
Qy 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTNTVILTNPISKEKQSHVSVLVYVP 420  
Db 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTNTVILTNPISKEKQSHVSVLVYVP 420  
Qy 421 PQIGKSLSPVDSYQVGTTLTCTVYAI PPPHHIHWYQWLEECANESQAVSVTNPY 480  
Db 421 PQIGKSLSPVDSYQVGTTLTCTVYAI PPPHHIHWYQWLEECANESQAVSVTNPY 480  
Qy 481 PCMEWRSVEDFQGNKI EYVKNQFALIEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE 540  
Db 481 PCMEWRSVEDFQGNKI EYVKNQFALIEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE 540  
Qy 541 RVISFHVTRGPEITLQPDMPQTEQESVSLWCTADRSTFENLTWYKLGPOPLPIHVGELPT 600  
Db 541 RVISFHVTRGPEITLQPDMPQTEQESVSLWCTADRSTFENLTWYKLGPOPLPIHVGELPT 600  
Qy 601 PUCKNLDTLWKNATWFSNSTNDILIMELKNASLQDQGVVCLAQDRKTKRHCVVROLT 660  
Db 601 PUCKNLDTLWKNATWFSNSTNDILIMELKNASLQDQGVVCLAQDRKTKRHCVVROLT 660  
Qy 661 VLERVAPTITGNLENQTTISGESIEVSTASGNPPPIIMWFKNDETILVEDSGIVLKDGNR 720  
Db 661 VLERVAPTITGNLENQTTISGESIEVSTASGNPPPIIMWFKNDETILVEDSGIVLKDGNR 720  
Qy 721 NLTI RVRKEDGLYTCQACSVLGC AKVBAFFIIEGAQKTNLEIILVGTAVIAMFFWL 780  
Db 721 NLTI RVRKEDGLYTCQACSVLGC AKVBAFFIIEGAQKTNLEIILVGTAVIAMFFWL 780  
Qy 781 LVLIIILTVKRGANGELKTYLSIVMDPDELPLDEHCERLPYDASKWEPRDRKLKGKPL 840  
Db 781 LVLIIILTVKRGANGELKTYLSIVMDPDELPLDEHCERLPYDASKWEPRDRKLKGKPL 840  
Qy 841 GRGAFQVIEADAFGIDKTATCTRTAVAKMLKEGATHSEHRLMSELKILIHIGHILNVN 900  
Db 841 GRGAFQVIEADAFGIDKTATCTRTAVAKMLKEGATHSEHRLMSELKILIHIGHILNVN 900  
Qy 901 LIGACTKPGGMLMVIIEFCKFNGLSYLRKSRNEFVYKTKGARFRQGDYVGAIPVDLK 960  
Db 901 LIGACTKPGGMLMVIIEFCKFNGLSYLRKSRNEFVYKTKGARFRQGDYVGAIPVDLK 960  
Qy 961 RRLDSITSSQSSASSGFVEEKSLSLSDVEEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
Db 961 RRLDSITSSQSSASSGFVEEKSLSLSDVEEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
Qy 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDDYVRKGDARPLKWMAPETIFDR 1080  
Db 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDDYVRKGDARPLKWMAPETIFDR 1080  
Qy 1081 VYTIQSDWMSFGVLLWEIFSLGASPYPGVKIDEEFCRLKEGTRMRAPDYTTPEMYQTML 1140  
Db 1081 VYTIQSDWMSFGVLLWEIFSLGASPYPGVKIDEEFCRLKEGTRMRAPDYTTPEMYQTML 1140  
Qy 1141 DCWHGEPSPRTFSELVEHLGNLLQANAQDGKDIYVLPISSETLSMEEDSGLSPTSPTS 1200  
Db 1141 DCWHGEPSPRTFSELVEHLGNLLQANAQDGKDIYVLPISSETLSMEEDSGLSPTSPTS 1200  
Qy 1201 CMEEEVCDPKFHYDNTAGISQYLQNSKRKSRPVSVKTFEDIPLEPEVKVIPPDDNQDTS 1260  
Db 1201 CMEEEVCDPKFHYDNTAGISQYLQNSKRKSRPVSVKTFEDIPLEPEVKVIPPDDNQDTS 1260  
Qy 1261 GMLVASEELKTLEDRTKLSFSGMVPSKRSRESVASEGNSQTSQYQSGYHSDDDTDTTVYS 1320  
Db 1261 GMLVASEELKTLEDRTKLSFSGMVPSKRSRESVASEGNSQTSQYQSGYHSDDDTDTTVYS 1320  
Qy 1321 SEAEALLKLIETGVGTGTAQILQPDGSGTTLSSPPV 1356  
Db 1321 SEAEALLKLIETGVGTGTAQILQPDGSGTTLSSPPV 1356

## RESULT 5

US-10-090-183-2  
; Sequence 2, Application US/10090183  
; Publication No. US20030185802A1  
; GENERAL INFORMATION:  
; APPLICANT: The Scripps Research Institute  
; APPLICANT: Ralph A. Reisfeld  
; APPLICANT: Andrew G. Niethammer  
; APPLICANT: Rong Xiang  
; TITLE OF INVENTION: DNA VACCINE AGAINST PROLIFERATING  
; TITLE OF INVENTION: ENDOTHELIAL CELLS AND METHODS OF USE THEREOF  
; FILE REFERENCE: TSRI-829.0  
; CURRENT APPLICATION NUMBER: US/10/090.183  
; CURRENT FILING DATE: 2002-03-02  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1356  
; TYPE: PRT  
; ORGANISM: human  
US-10-090-183-2

Query Match 99.9%; Score 7092; DB 4; Length 1356;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MESKVLAVLWLCVETRAASVGLPSVSLDPLRLSIQKDIITIKANTTLQITCRGORDLD	60
Db	1	MOSKVLAVLWLCVETRAASVGLPSVSLDPLRLSIQKDIITIKANTTLQITCRGORDLD	60
Qy	61	WLPNNQSGEORVEVEVTECSGDLFCCKTLTIKPVIGNDTGAYKCFYRETDLASVIVYVQD	120
Db	61	WLPNNQSGEORVEVEVTECSGDLFCCKTLTIKPVIGNDTGAYKCFYRETDLASVIVYVQD	120
Qy	121	YRSPFIASVSDQGVVYITENKNTVPICLGSIINLVNLSLCARYPEKRFVDPGNRISWD	180
Db	121	YRSPFIASVSDQGVVYITENKNTVPICLGSIINLVNLSLCARYPEKRFVDPGNRISWD	180
Qy	181	SKGFTIPSMISVAGMVFCEAKINDESYOSIMYIVVVVGYRIYDVVLSPSHGIELSVGE	240
Db	181	SKGFTIPSMISVAGMVFCEAKINDESYOSIMYIVVVVGYRIYDVVLSPSHGIELSVGE	240
Qy	241	KLVLNCTARTALNVGIDFNWEYSSKHOKKLVNRDLKTGSGSEMCKFLSTLIDGVTRS	300
Db	241	KLVLNCTARTALNVGIDFNWEYSSKHOKKLVNRDLKTGSGSEMCKFLSTLIDGVTRS	300
Qy	301	DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVIRIPAKYLGYPVP	360
Db	301	DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVIRIPAKYLGYPVP	360
Qy	361	EIKWYKNGIPILESNHTIKAGHVLTIMEVSRDGTGNVTILNTPISKEKQSHVSVLVYVP	420
Db	361	EIKWYKNGIPILESNHTIKAGHVLTIMEVSRDGTGNVTILNTPISKEKQSHVSVLVYVP	420
Qy	421	POIGKSLIPSVDSYQGTITCTVYAIIPPHIHWHYQWLEECANESQAVSTNYP	480
Db	421	POIGKSLIPSVDSYQGTITCTVYAIIPPHIHWHYQWLEECANESQAVSTNYP	480
Qy	481	PCBEWSRVEDFOGKNLEVNKNQFALLTEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE	540
Db	481	PCBEWSRVEDFOGKNLEVNKNQFALLTEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE	540
Qy	541	RVISFHVTRGPETITLOPDMQTEQESVSLWCTADRSTFENLTWYKLPQPLIHVGELPT	600
Db	541	RVISFHVTRGPETITLOPDMQTEQESVSLWCTADRSTFENLTWYKLPQPLIHVGELPT	600
Qy	601	PVCNKLDTLWKLNATMFNSNTDILIMELKNASLODQGDYVCLAOADRKTKKRCHVVRQLT	660
Db	601	PVCNKLDTLWKLNATMFNSNTDILIMELKNASLODQGDYVCLAOADRKTKKRCHVVRQLT	660
Qy	661	VLBRVAPITIGNLENQTTISGESIEVSTASGPNPPQIMWFKNETIIVEDSGIVLKDGNR	720
Db	661	VLBRVAPITIGNLENQTTISGESIEVSTASGPNPPQIMWFKNETIIVEDSGIVLKDGNR	720

Qy	721	NLTIRVRKEDEGLYTCQACSVLGCACVFAFFIIIEGAQKTNLEIIILVCTAVIAMEFWL	780
Db	721	NLTIRVRKEDEGLYTCQACSVLGCACVFAFFIIIEGAQKTNLEIIILVCTAVIAMEFWL	780
Qy	781	LLVILRTVKRANGGELKGTGYSIVMDPDELPLDEHCERLPYDASKWEFFPRDRILKLGKPL	840
Db	781	LLVILRTVKRANGGELKGTGYSIVMDPDELPLDEHCERLPYDASKWEFFPRDRILKLGKPL	840
Qy	841	GRGAFQVTEADAFGIDKTATCTVAVKMLKEGATHSEHRLMSELKILIHGHLLNVN	900
Db	841	GRGAFQVTEADAFGIDKTATCTVAVKMLKEGATHSEHRLMSELKILIHGHLLNVN	900
Qy	901	LLGACTKPGGLMVIYVEFCCKFNLSTVLSRKNEFVPIYKTKGARFRGQKDYVGAIPVDLK	960
Db	901	LLGACTKPGGLMVIYVEFCCKFNLSTVLSRKNEFVPIYKTKGARFRGQKDYVGAIPVDLK	960
Qy	961	RRLDSITSSQSSASSGSGFVEEKSLSVDEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA	1020
Db	961	RRLDSITSSQSSASSGSGFVEEKSLSVDEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA	1020
Qy	1021	SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR	1080
Db	1021	SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR	1080
Qy	1081	VYTIQSDVMSFGVLLWEIIFSLGASPYGVKIDEEFCERLKEGTRMRAPDYTTPEMYOTML	1140
Db	1081	VYTIQSDVMSFGVLLWEIIFSLGASPYGVKIDEEFCERLKEGTRMRAPDYTTPEMYOTML	1140
Qy	1141	DCWHGSPSORPTSELSVEHLNLLQANAQODGKDYIVLPISETLSMEEDSGLSPTSPTS	1200
Db	1141	DCWHGSPSORPTSELSVEHLNLLQANAQODGKDYIVLPISETLSMEEDSGLSPTSPTS	1200
Qy	1201	CMEEBVCDDPKHYDNTAGISQYLNQSKRSRPSVKTFFEDIPLEBPEVKVIPPDDNQDTS	1260
Db	1201	CMEEBVCDDPKHYDNTAGISQYLNQSKRSRPSVKTFFEDIPLEBPEVKVIPPDDNQDTS	1260
Qy	1261	GMVLASEELKLEIGVQGTSTAQIILQPDGTTLSPPV	1320
Db	1261	GMVLASEELKLEIGVQGTSTAQIILQPDGTTLSPPV	1320
Qy	1321	SEBAELKLEIGVQGTSTAQIILQPDGTTLSPPV	1356
Db	1321	SEBAELKLEIGVQGTSTAQIILQPDGTTLSPPV	1356

## RESULT 6

US-10-394-322A-66  
; Sequence 66, Application US/10394322A  
; Publication No. US20030232391A1  
; GENERAL INFORMATION:  
; APPLICANT: SUNESIS PHARMACEUTICALS, INC.  
; APPLICANT: Prescott, John C.  
; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS  
; FILE REFERENCE: 39750-0006 US  
; CURRENT APPLICATION NUMBER: US/10/394,322A  
; PRIOR FILING DATE: 2003-03-20  
; PRIOR FILING DATE: 2002-03-21  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 66  
; LENGTH: 1356  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-394-322A-66

Query Match 99.9%; Score 7092; DB 4; Length 1356;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MESKVLAVLWLCVETRAASVGLPSVSLDPLRLSIQKDIITIKANTTLQITCRGORDLD	60
Db	1	MESKVLAVLWLCVETRAASVGLPSVSLDPLRLSIQKDIITIKANTTLQITCRGORDLD	60

Db	1	MQSKVLLAVALWLCVETRAASVGLPSVSLDPLRLSIQKDLITIKANTTLQITCRGQDRLD	60
Qy	61	WLWPNNSGSEQRVEVTECSGDLFCCKTLTI PKVIGNDTGAYKCFYRETDLASVIVYVQD	120
Db	61	WLWPNNSGSEQRVEVTECSGDLFCCKTLTI PKVIGNDTGAYKCFYRETDLASVIVYVQD	120
Qy	121	YRSPFIASVDQHGVIYITENKNTVVI PCLGSI SNLNSLCARYPKRFPVDPGNRISWD	180
Db	121	YRSPFIASVDQHGVIYITENKNTVVI PCLGSI SNLNSLCARYPKRFPVDPGNRISWD	180
Qy	181	SKGFTIPSMYISYAGWVFCEAKINDESYQIMYIVVVVGYRIYDVVLSPSHGIELSVGE	240
Db	181	SKGFTIPSMYISYAGWVFCEAKINDESYQIMYIVVVVGYRIYDVVLSPSHGIELSVGE	240
Qy	241	KLVLNCTARTELNVGIDFNWYPSSSHQHKLVNRDLKTQSGSEMCKFLSTLTIDGVTRS	300
Db	241	KLVLNCTARTELNVGIDFNWYPSSSHQHKLVNRDLKTQSGSEMCKFLSTLTIDGVTRS	300
Qy	301	DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVIRIPAKYLYGPPP	360
Db	301	DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVIRIPAKYLYGPPP	360
Qy	361	EIKWYKNGIPLSNHTIKAGHVLTIMEVSRDGTGNTVILTNPISKEQSHVSVLVVYVP	420
Db	361	EIKWYKNGIPLSNHTIKAGHVLTIMEVSRDGTGNTVILTNPISKEQSHVSVLVVYVP	420
Qy	421	POIGEKSLISPVDSYQGTGTTCTVYALPPPHIHWYQLEBEECANEPQAVSVTNPY	480
Db	421	POIGEKSLISPVDSYQGTGTTCTVYALPPPHIHWYQLEBEECANEPQAVSVTNPY	480
Qy	481	PCBEWRSVEDFOGNGKIEVKNQFALIEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE	540
Db	481	PCBEWRSVEDFOGNGKIEVKNQFALIEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE	540
Qy	541	RVISFHVTRGPETITLOPDMQTOESVSLWCTADRSTFENLTWYKGPQLPIHVGEELPT	600
Db	541	RVISFHVTRGPETITLOPDMQTOESVSLWCTADRSTFENLTWYKGPQLPIHVGEELPT	600
Qy	601	PVCNKDLTLWKNATMFSNSTDILIMELKNASIQODGDVVCIAQDRKTKRHCVVRLT	660
Db	601	PVCNKDLTLWKNATMFSNSTDILIMELKNASIQODGDVVCIAQDRKTKRHCVVRLT	660
Qy	661	VLERVAPTIITGNLENQTTISGESIEVSCASGNPPQIMWFKDNETLVEDSGIVLKDGNR	720
Db	661	VLERVAPTIITGNLENQTTISGESIEVSCASGNPPQIMWFKDNETLVEDSGIVLKDGNR	720
Qy	721	NLTIRVRKEDEGLYTQCAQSVLGCACVVEAFPIIEGAQEKTNLEIILVGTAVIAMFFWL	780
Db	721	NLTIRVRKEDEGLYTQCAQSVLGCACVVEAFPIIEGAQEKTNLEIILVGTAVIAMFFWL	780
Qy	781	LLVILIRTVKRVANGELKTYLSIVMDPDELPLDEHCERLPYDASKWEPPDRRLKLGKPL	840
Db	781	LLVILIRTVKRVANGELKTYLSIVMDPDELPLDEHCERLPYDASKWEPPDRRLKLGKPL	840
Qy	841	GRGAFGOVIEADAFGIDKATCTRTVAVKMLKEGATHSEHRALMSELKILIHIGHLNVVN	900
Db	841	GRGAFGOVIEADAFGIDKATCTRTVAVKMLKEGATHSEHRALMSELKILIHIGHLNVVN	900
Qy	901	LLGACTKPGGPLMVI VEFCKFGNLSYLRSKRNEFPVYTKGARFRQGDYGAIPVDLK	960
Db	901	LLGACTKPGGPLMVI VEFCKFGNLSYLRSKRNEFPVYTKGARFRQGDYGAIPVDLK	960
Qy	961	RLRDSITSSOSSASSGFVEEKSISDVEEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA	1020
Db	961	RLRDSITSSOSSASSGFVEEKSISDVEEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA	1020
Qy	1021	SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWWAPETIFDR	1080
Db	1021	SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWWAPETIFDR	1080
Qy	1081	VYTIQSDVNSFGVLLWEIFSLGASPYGVKIDBEFCRRLKEGTRMRAPDYTTPEMYQTML	1140
Db	1081	VYTIQSDVNSFGVLLWEIFSLGASPYGVKIDBEFCRRLKEGTRMRAPDYTTPEMYQTML	1140

Qy	1141	DCWHGSPSORPTSELSVEHLGNLLQANAQODGKDYIVLPISETLSMEEDSGLSLPTSPVS	1200
Db	1141	DCWHGSPSORPTSELSVEHLGNLLQANAQODGKDYIVLPISETLSMEEDSGLSLPTSPVS	1200
Qy	1201	CMEEEVCDPKFHYDNTAGISQYQLNSKRKRSPVSVKTFEDI PLEBPEVKVIBDDNQDTS	1260
Db	1201	CMEEEVCDPKFHYDNTAGISQYQLNSKRKRSPVSVKTFEDI PLEBPEVKVIBDDNQDTS	1260
Qy	1261	GMVLASEELKLTEDRKTLSPSFGMVPSKSRRESVASEGNSQTSYQSGYHSDDDTDTTVYS	1320
Db	1261	GMVLASEELKLTEDRKTLSPSFGMVPSKSRRESVASEGNSQTSYQSGYHSDDDTDTTVYS	1320
Qy	1321	SEEAELLKLTIEIGVQTGSTAQIQLQPDSTGTTLSPPV	1356
Db	1321	SEEAELLKLTIEIGVQTGSTAQIQLQPDSTGTTLSPPV	1356
RESULT 7			
US-10-440-464-129			
; Sequence 129, Application US/10440464			
; Publication No. US20040018528A1			
; GENERAL INFORMATION:			
; APPLICANT: DEPRIMO, SAMUEL			
; APPLICANT: O'FARRELL, ANNE-MARIE			
; APPLICANT: MORIMOTO, ALYSSA			
; APPLICANT: SMOLICH, BEVERLY			
; APPLICANT: MANNING, WILLIAM			
; APPLICANT: WALTER, SARAH			
; APPLICANT: CERRINGTON, JULIE			
; APPLICANT: SCHILLING, JIM			
; TITLE OF INVENTION: NOVEL BIOMARKERS OF TYROSINE KINASE INHIBITOR EXPOSURE			
; TITLE OF INVENTION: AND ACTIVITY IN MAMMALS			
; FILE REFERENCE: 038602/1592			
; CURRENT APPLICATION NUMBER: US/10/440,464			
; PRIOR FILING DATE: 2003-05-19			
; PRIOR APPLICATION NUMBER: 60/380,872			
; PRIOR FILING DATE: 2002-05-17			
; PRIOR APPLICATION NUMBER: 60/448,922			
; PRIOR FILING DATE: 2003-02-24			
; PRIOR APPLICATION NUMBER: 60/448,874			
; PRIOR FILING DATE: 2003-02-24			
; NUMBER OF SEQ ID NOS: 185			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 129			
; LENGTH: 1356			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-440-464-129			
Query Match 99.9%; Score 7092; DB 4; Length 1356;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MBSKVLLAVALWLCVETRAASVGLPSVSLDPLRLSIQKDLITIKANTTLQITCRGQDRLD	60
Db	1	MBSKVLLAVALWLCVETRAASVGLPSVSLDPLRLSIQKDLITIKANTTLQITCRGQDRLD	60
Qy	61	WLWPNNSGSEQRVEVTECSGDLFCCKTLTI PKVIGNDTGAYKCFYRETDLASVIVYVQD	120
Db	61	WLWPNNSGSEQRVEVTECSGDLFCCKTLTI PKVIGNDTGAYKCFYRETDLASVIVYVQD	120
Qy	121	YRSPFIASVDQHGVIYITENKNTVVI PCLGSI SNLNSLCARYPKRFPVDPGNRISWD	180
Db	121	YRSPFIASVDQHGVIYITENKNTVVI PCLGSI SNLNSLCARYPKRFPVDPGNRISWD	180
Qy	181	SKGFTIPSMYISYAGWVFCEAKINDESYQIMYIVVVVGYRIYDVVLSPSHGIELSVGE	240
Db	181	SKGFTIPSMYISYAGWVFCEAKINDESYQIMYIVVVVGYRIYDVVLSPSHGIELSVGE	240
Qy	241	KLVLNCTARTELNVGIDFNWYPSSSHQHKLVNRDLKTQSGSEMCKFLSTLTIDGVTRS	300
Db	241	KLVLNCTARTELNVGIDFNWYPSSSHQHKLVNRDLKTQSGSEMCKFLSTLTIDGVTRS	300



```
Qy 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIIPAKYLGYPVP 360
Db 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIIPAKYLGYPVP 360
Qy 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTGNTVILTNPISEKQSHVSVLVVYP 420
Db 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTGNTVILTNPISEKQSHVSVLVVYP 420
Qy 421 PQIGEXSLSPVDSYQYGTQTTLCTVYAIIPPHHHIHWYQLEBECANESQAVSVTNPY 480
Db 421 PQIGEXSLSPVDSYQYGTQTTLCTVYAIIPPHHHIHWYQLEBECANESQAVSVTNPY 480
Qy 481 PCBEWRSVEDFOGKNKIEVNKNQFALIEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE 540
Db 481 PCBEWRSVEDFOGKNKIEVNKNQFALIEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE 540
Qy 541 RVISFHVTRGPETITLQDMQPTQESVSLWCTADRSTFENLTWYKLGQPLPIHVGELPT 600
Db 541 RVISFHVTRGPETITLQDMQPTQESVSLWCTADRSTFENLTWYKLGQPLPIHVGELPT 600
Qy 601 PUCKNLDTLKLNATMFSNSTNDILIMELKNASLQDQGVYCLAQDRKTKKRHCVRQLT 660
Db 601 PUCKNLDTLKLNATMFSNSTNDILIMELKNASLQDQGVYCLAQDRKTKKRHCVRQLT 660
Qy 661 VLERVAPTTIGNLENOTTISIGESIEVSCASGNPPPOIMWFKNETLVEDSGIVLKDGNR 720
Db 661 VLERVAPTTIGNLENOTTISIGESIEVSCASGNPPPOIMWFKNETLVEDSGIVLKDGNR 720
Qy 721 NLITRRVKEDEGLYTQCAQSVLGCACVAFIEGAEQKTNLEIILVGTAVIAMPFWL 780
Db 721 NLITRRVKEDEGLYTQCAQSVLGCACVAFIEGAEQKTNLEIILVGTAVIAMPFWL 780
Qy 781 LLVILITVRKANGELKTYLSVMDPDELPLEHCEKRLPYDASKWEPDRDLKLGKPL 840
Db 781 LLVILITVRKANGELKTYLSVMDPDELPLEHCEKRLPYDASKWEPDRDLKLGKPL 840
Qy 841 GRGAFQGVIEADAFIDKATCRTVAVKMLKEGATHSEHRLMSELKILIHGHILNVN 900
Db 841 GRGAFQGVIEADAFIDKATCRTVAVKMLKEGATHSEHRLMSELKILIHGHILNVN 900
Qy 901 LLGACTKPGGLMVIIEFCFGLNLTSLRKRNEFVYKTKGARFROGKYVGAIPVDLK 960
Db 901 LLGACTKPGGLMVIIEFCFGLNLTSLRKRNEFVYKTKGARFROGKYVGAIPVDLK 960
Qy 961 RRLDSITSSQSSASSGFVEBSKLSVDEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020
Db 961 RRLDSITSSQSSASSGFVEBSKLSVDEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020
Qy 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDDPYVRKGDARLPLKWMAPETIFDR 1080
Db 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDDPYVRKGDARLPLKWMAPETIFDR 1080
Qy 1081 VYTIQSDVWFGVLLWEIFSLGASPYGVKIDEEFCRLKEGTRMRAPDYTTPEMYQTM 1140
Db 1081 VYTIQSDVWFGVLLWEIFSLGASPYGVKIDEEFCRLKEGTRMRAPDYTTPEMYQTM 1140
Qy 1141 DCWHGEPSPRTFSELVEHLGNLLQANAQDQGDYIIVLPISSETLSMEEDSGLSPTSPVS 1200
Db 1141 DCWHGEPSPRTFSELVEHLGNLLQANAQDQGDYIIVLPISSETLSMEEDSGLSPTSPVS 1200
Qy 1201 CMEEBEEVCDPKFHYDNTAGISQYLQNSKRKSRPVSVKTFPEDIPLBEPVVKVIPPDDNQDTS 1260
Db 1201 CMEEBEEVCDPKFHYDNTAGISQYLQNSKRKSRPVSVKTFPEDIPLBEPVVKVIPPDDNQDTS 1260
Qy 1261 GMLVASELKTLEDRTKLSFGCGWVPSKRESVASEGNSQTSQYQSGYHSDDDTDTTVYS 1320
Db 1261 GMLVASELKTLEDRTKLSFGCGWVPSKRESVASEGNSQTSQYQSGYHSDDDTDTTVYS 1320
Qy 1321 SEEAELKLIEIGVQTGSTAQIQLQPDGTTLSGPPV 1356
Db 1321 SEEAELKLIEIGVQTGSTAQIQLQPDGTTLSGPPV 1356
```

```
RESULT 8
US-10-783-528-61
; Sequence 61, Application US/10783528
; Publication No. US20040219579A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Gish, Kurt
; APPLICANT: Wilson, Keith
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: METHODS OF DIAGNOSIS OF CANCER, COMPOSITIONS AND
; FILE REFERENCE: 05882.0191.NPUS01
; CURRENT APPLICATION NUMBER: US/10/783,528
; CURRENT FILING DATE: 2004-02-19
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 61
; LENGTH: 1356
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-783-528-61
```

```
Query Match 99.9%; Score 7092; DB 5; Length 1356;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MESKULLAVALMILCVETRAASVGLPSVSLDPLRLSQDKILTIKANTTLOITCRGQRDL 60
Db 1 MQSKULLAVALMILCVETRAASVGLPSVSLDPLRLSQDKILTIKANTTLOITCRGQRDL 60
Qy 61 WLPNNQSGSEQRVEVTECSGDLFCCKTLTIPIKVINGTGAYKCFYRTDLASVIVVQD 120
Db 61 WLPNNQSGSEQRVEVTECSGDLFCCKTLTIPIKVINGTGAYKCFYRTDLASVIVVQD 120
Qy 121 YRSPFIASVSDQGVVYITENKNTVYIPCLGSIINLNSLCARYPEKRFVDPGNRISWD 180
Db 121 YRSPFIASVSDQGVVYITENKNTVYIPCLGSIINLNSLCARYPEKRFVDPGNRISWD 180
Qy 181 SKKGFTIPSMIYAGWVCEAKINDESYQIMYIVVWGYRYIDVVLSPSHGIELSVGE 240
Db 181 SKKGFTIPSMIYAGWVCEAKINDESYQIMYIVVWGYRYIDVVLSPSHGIELSVGE 240
Qy 241 KLVLNCTARTELNVGIDENWEYPSKXHKHKLNRDLKTQSGSEMKKFLTLTIDGVT 300
Db 241 KLVLNCTARTELNVGIDENWEYPSKXHKHKLNRDLKTQSGSEMKKFLTLTIDGVT 300
Qy 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIIPAKYLGYPVP 360
Db 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIIPAKYLGYPVP 360
Qy 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTGNTVILTNPISEKQSHVSVLVVYP 420
Db 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTGNTVILTNPISEKQSHVSVLVVYP 420
Qy 421 PQIGEXSLSPVDSYQYGTQTTLCTVYAIIPPHHHIHWYQLEBECANESQAVSVTNPY 480
Db 421 PQIGEXSLSPVDSYQYGTQTTLCTVYAIIPPHHHIHWYQLEBECANESQAVSVTNPY 480
Qy 481 PCBEWRSVEDFOGKNKIEVNKNQFALIEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE 540
Db 481 PCBEWRSVEDFOGKNKIEVNKNQFALIEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE 540
Qy 541 RVISFHVTRGPETITLQDMQPTQESVSLWCTADRSTFENLTWYKLGQPLPIHVGELPT 600
Db 541 RVISFHVTRGPETITLQDMQPTQESVSLWCTADRSTFENLTWYKLGQPLPIHVGELPT 600
Qy 601 PUCKNLDTLKLNATMFSNSTNDILIMELKNASLQDQGVYCLAQDRKTKKRHCVRQLT 660
Db 601 PUCKNLDTLKLNATMFSNSTNDILIMELKNASLQDQGVYCLAQDRKTKKRHCVRQLT 660
Qy 661 VLERVAPTTIGNLENOTTISIGESIEVSCASGNPPPOIMWFKNETLVEDSGIVLKDGNR 720
Db 661 VLERVAPTTIGNLENOTTISIGESIEVSCASGNPPPOIMWFKNETLVEDSGIVLKDGNR 720
```

721 NLTIIRVRKEDGLYTCQACSVLGCACVAFPIIEGAQEKTNLEIILVGTAVIAMFFWL 780  
721 NLTIIRVRKEDGLYTCQACSVLGCACVAFPIIEGAQEKTNLEIILVGTAVIAMFFWL 780  
781 LLVILITVRANGGELKTYLSVMDPDELDEHCERLPYDASKWEFFPRDRKLKPKL 840  
781 LLVILITVRANGGELKTYLSVMDPDELDEHCERLPYDASKWEFFPRDRKLKPKL 840  
841 GRGAFQGVIEADAFIDKATCTRTVAVMLKEGATHSEHRLMSELKILIHGHILNVN 900  
841 GRGAFQGVIEADAFIDKATCTRTVAVMLKEGATHSEHRLMSELKILIHGHILNVN 900  
901 LLGACTKPGGLMVIYVEFCFKNLSTYLSRKNEFPYKTKGARFRQGDYVGAIPVDLK 960  
901 LLGACTKPGGLMVIYVEFCFKNLSTYLSRKNEFPYKTKGARFRQGDYVGAIPVDLK 960  
961 RLDSITSSQSSASSGFVEBSLSVDEEAPEDLYKDFLTLEHLICYSPQVAKMEFLA 1020  
961 RLDSITSSQSSASSGFVEBSLSVDEEAPEDLYKDFLTLEHLICYSPQVAKMEFLA 1020  
1021 SRKCIHRDLAARNILISEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR 1080  
1021 SRKCIHRDLAARNILISEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR 1080  
1081 VYTIQSDVMSFGVLLWEIFSLGASPYGVKIDBEFCRRKKEGTRMRAPDYTTPEMYQTM 1140  
1081 VYTIQSDVMSFGVLLWEIFSLGASPYGVKIDBEFCRRKKEGTRMRAPDYTTPEMYQTM 1140  
1141 DCWHGSPQSPPTSELSVEHLGNLQANAQODGKDYIVLPISETLSMEEDSGLSPTSPVS 1200  
1141 DCWHGSPQSPPTSELSVEHLGNLQANAQODGKDYIVLPISETLSMEEDSGLSPTSPVS 1200  
1201 CMEEEVCDPKFYDNTAGISQYLQNSKRKSRPVSKVTEDIDLEBPVKVIPPDDNQTDS 1260  
1201 CMEEEVCDPKFYDNTAGISQYLQNSKRKSRPVSKVTEDIDLEBPVKVIPPDDNQTDS 1260  
1261 GVLVASEELKLEIGVQTGSTAQIILQPSGTTLSPPV 1356  
1261 GVLVASEELKLEIGVQTGSTAQIILQPSGTTLSPPV 1356  
1321 SEBAELLKLEIGVQTGSTAQIILQPSGTTLSPPV 1356  
1321 SEBAELLKLEIGVQTGSTAQIILQPSGTTLSPPV 1356

RESULT 9

US-10-872-198-115  
; Sequence 115, Application US/10872198  
; Publication No. US20050002897A1  
; GENERAL INFORMATION:  
; APPLICANT: Ulrich HAUPT  
; APPLICANT: Andre KOLTERMANN  
; APPLICANT: Andreas SCHEIDIG  
; APPLICANT: Christian VOETSMEIER  
; APPLICANT: Ulrich Ketting  
; TITLE OF INVENTION: NEW BIOLOGICAL ENTITIES AND USE THEREOF  
; FILE REFERENCE: 04156.000204  
; CURRENT APPLICATION NUMBER: US/10/872,198  
; CURRENT FILING DATE: 2004-06-18  
; PRIOR APPLICATION NUMBER: 60/543,518  
; PRIOR FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: 60/524,960  
; PRIOR FILING DATE: 2003-11-25  
; PRIOR APPLICATION NUMBER: EP 04003058  
; PRIOR FILING DATE: 2004-02-11  
; PRIOR APPLICATION NUMBER: EP 03025871  
; PRIOR FILING DATE: 2003-11-11  
; PRIOR APPLICATION NUMBER: EP 03025851  
; PRIOR FILING DATE: 2003-11-10  
; PRIOR APPLICATION NUMBER: EP 03013819  
; PRIOR FILING DATE: 2003-06-18  
; NUMBER OF SEQ ID NOS: 149

; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 115  
; LENGTH: 1356  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-872-198-115  
  
Query Match 99.9%; Score 7092; DB 5; Length 1356;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MESKVLALVALMVCVETRAASVGLPSVSLDLPRLSIQKIDILTIKANTTLOITCRGORDLD 60  
DB 1 MOSKVLALVALMVCVETRAASVGLPSVSLDLPRLSIQKIDILTIKANTTLOITCRGORDLD 60  
QY 61 WLPNQQSSSEORVEVTECSGDLFCCKTLTIPKVIKNDTGAYKCFYRETDLASVYVYVQD 120  
DB 61 WLPNQQSSSEORVEVTECSGDLFCCKTLTIPKVIKNDTGAYKCFYRETDLASVYVYVQD 120  
QY 121 YRSPFTIASVSDQHGYYVYITENKNTVVIICLGSISNLNLSLCARYPEKRFVDPGNRISWD 180  
DB 121 YRSPFTIASVSDQHGYYVYITENKNTVVIICLGSISNLNLSLCARYPEKRFVDPGNRISWD 180  
QY 181 SKGFTIPSMISYAGWVCEAKINDESYQSIYIVVVYGYRIYDVVLSHGHIELSVGE 240  
DB 181 SKGFTIPSMISYAGWVCEAKINDESYQSIYIVVVYGYRIYDVVLSHGHIELSVGE 240  
QY 241 KLVNLCARTELNVGIDFNWEYPSKQHKLVNRLDKTQSGSEMKEFSTLTIDGVTRS 300  
DB 241 KLVNLCARTELNVGIDFNWEYPSKQHKLVNRLDKTQSGSEMKEFSTLTIDGVTRS 300  
QY 301 DOGLYTCASSGLMTKKNSTFVRVHEKPFVAFSGSMESLVEATVGERVRIPAKYLGYPPP 360  
DB 301 DOGLYTCASSGLMTKKNSTFVRVHEKPFVAFSGSMESLVEATVGERVRIPAKYLGYPPP 360  
QY 361 EIKWYKNGIPLSNHTIKAGHVLTINVESESDTGNVTILTNPISEKQSHVSVLVVYVP 420  
DB 361 EIKWYKNGIPLSNHTIKAGHVLTINVESESDTGNVTILTNPISEKQSHVSVLVVYVP 420  
QY 421 PQIGKSLISPVDSYQYGTQTLTCTVYAI PPHHIIHWYQLEEECANPSQAVSTNPNY 480  
DB 421 PQIGKSLISPVDSYQYGTQTLTCTVYAI PPHHIIHWYQLEEECANPSQAVSTNPNY 480  
QY 481 PCEEWRSEVEDFOGKNKIEVNKNQFALIEGKNKTSTLVIQAANVSALYKCEAVNKVGRGE 540  
DB 481 PCEEWRSEVEDFOGKNKIEVNKNQFALIEGKNKTSTLVIQAANVSALYKCEAVNKVGRGE 540  
QY 541 RVISPHVTRGPETITLQPMQTEQSVSLWCTADRTSTFENLTWYKLGPOPLPIHVGELPT 600  
DB 541 RVISPHVTRGPETITLQPMQTEQSVSLWCTADRTSTFENLTWYKLGPOPLPIHVGELPT 600  
QY 601 PVCKNLDTLWKLNAWFSNSTNDILIMELKNASLOQDGYVCLADRKTKRHCVVRLT 660  
DB 601 PVCKNLDTLWKLNAWFSNSTNDILIMELKNASLOQDGYVCLADRKTKRHCVVRLT 660  
QY 661 VLERVAPTTITGNLENQTTISGESIEVSCASGNPPPIIMFKDNETLVEDSGIVLKDGNR 720  
DB 661 VLERVAPTTITGNLENQTTISGESIEVSCASGNPPPIIMFKDNETLVEDSGIVLKDGNR 720  
QY 721 NLTIRVRKEDGLYTCQACSVLGCACVAFPIIEGAQEKTNLEIILVGTAVIAMFFWL 780  
DB 721 NLTIRVRKEDGLYTCQACSVLGCACVAFPIIEGAQEKTNLEIILVGTAVIAMFFWL 780  
QY 781 LLVILITVRANGGELKTYLSVMDPDELDEHCERLPYDASKWEFFPRDRKLKPKL 840  
DB 781 LLVILITVRANGGELKTYLSVMDPDELDEHCERLPYDASKWEFFPRDRKLKPKL 840  
QY 841 GRGAFQGVIEADAFIDKATCTRTVAVMLKEGATHSEHRLMSELKILIHGHILNVN 900  
DB 841 GRGAFQGVIEADAFIDKATCTRTVAVMLKEGATHSEHRLMSELKILIHGHILNVN 900  
QY 901 LLGACTKPGGLMVIYVEFCFKNLSTYLSRKNEFPYKTKGARFRQGDYVGAIPVDLK 960  
DB 901 LLGACTKPGGLMVIYVEFCFKNLSTYLSRKNEFPYKTKGARFRQGDYVGAIPVDLK 960

Db 901 LLGACTKPGGLMVI VEFCKFGNLSYLRSKRNEFVVKTKGARFRQGDYVGAIPVDLK 960  
Qy 961 RRLDSITSSQSSASSGFGVEEKSLSDVVEEBEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
Db 961 RRLDSITSSQSSASSGFGVEEKSLSDVVEEBEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
Qy 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPKWMAPETIFDR 1080  
Db 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPKWMAPETIFDR 1080  
Qy 1081 VYTIQSDWMSFGVLLWEIFSLGASPGVKIDEEFCRRLEKGTMRAPDYTTPEMYOTML 1140  
Db 1081 VYTIQSDWMSFGVLLWEIFSLGASPGVKIDEEFCRRLEKGTMRAPDYTTPEMYOTML 1140  
Qy 1141 DCWHGEPSPQPTSELSVEHLGNLLQANAQDGKDYIVLPISSETLSMEEDSGLSLPTSPVS 1200  
Db 1141 DCWHGEPSPQPTSELSVEHLGNLLQANAQDGKDYIVLPISSETLSMEEDSGLSLPTSPVS 1200  
Qy 1201 CMEEEVCDPKFHYDNTAGISQYLQNSKRKRSPVSVKTFEDIPLIEEPVKVIPPDDNOTDS 1260  
Db 1201 CMEEEVCDPKFHYDNTAGISQYLQNSKRKRSPVSVKTFEDIPLIEEPVKVIPPDDNOTDS 1260  
Qy 1261 GWVLASEELKLTLEDRTKLSPSFGGMPVSKSRVASESGNSQTSYQSGYHSDDDTTTVYS 1320  
Db 1261 GWVLASEELKLTLEDRTKLSPSFGGMPVSKSRVASESGNSQTSYQSGYHSDDDTTTVYS 1320  
Qy 1321 SEAEALLKLTLEIGVQTGSTAQILQPDGTTLSPPV 1356  
Db 1321 SEAEALLKLTLEIGVQTGSTAQILQPDGTTLSPPV 1356

## RESULT 10

US-10-741-600-1469  
; Sequence 1469, Application US/10741600  
; Publication No. US20050026169A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001499  
; CURRENT APPLICATION NUMBER: US/10/741.600  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 73997  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1469  
; LENGTH: 1356  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-741-600-1469

Query Match 99.9%; Score 7092; DB 5; Length 1356;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESKVLLAVALWLCVETRAASVGLPSVSLDPLRLSIQDKILTIKANTLTQITCRGQDLD 60  
Db 1 MQSKVLLAVALWLCVETRAASVGLPSVSLDPLRLSIQDKILTIKANTLTQITCRGQDLD 60  
Qy 61 WLWPNNGSGRVEVEVTECDGLFCKTLTIPKVLGNDTGAYKCFYRETDLASVIYVYVQD 120  
Db 61 WLWPNNGSGRVEVEVTECDGLFCKTLTIPKVLGNDTGAYKCFYRETDLASVIYVYVQD 120  
Qy 121 YRSFFIASVSDQHGVIYITENKNTVVIPLCGSISNLNLSVCARYPEKRFVDPGNRISWD 180  
Db 121 YRSFFIASVSDQHGVIYITENKNTVVIPLCGSISNLNLSVCARYPEKRFVDPGNRISWD 180  
Qy 181 SKGFTTIPSYMISVAGMVFCEAKINDESYQIMYIVVVGRIYDVVLSPSHGIELSVGE 240  
Db 181 SKGFTTIPSYMISVAGMVFCEAKINDESYQIMYIVVVGRIYDVVLSPSHGIELSVGE 240  
Qy 241 KVLNCTARTELNVGIDFNWEYPSKQHKLVNRDLKTQSGSMKKFLSTLTIDGVTRS 300  
Db 241 KVLNCTARTELNVGIDFNWEYPSKQHKLVNRDLKTQSGSMKKFLSTLTIDGVTRS 300

Qy 301 DOGLYTCAASGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVIRPAKYLGYPPP 360  
Db 301 DOGLYTCAASGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVIRPAKYLGYPPP 360  
Qy 361 EIKWYKNGIIPLESNHTIKAGHVLTIMESVSRDGTGYVTILTNPISEKQSHVSVLVVVP 420  
Db 361 EIKWYKNGIIPLESNHTIKAGHVLTIMESVSRDGTGYVTILTNPISEKQSHVSVLVVVP 420  
Qy 421 PQIGEKSLISPDVSYQYGTTLTCTVYAIIPPHHIIHWYQLEEBECANESQAVSVNTPY 480  
Db 421 PQIGEKSLISPDVSYQYGTTLTCTVYAIIPPHHIIHWYQLEEBECANESQAVSVNTPY 480  
Qy 481 PCBEWRSVEDFOGKNKIEVNKQFALIEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE 540  
Db 481 PCBEWRSVEDFOGKNKIEVNKQFALIEGKNKTVSTLVIQAAVNSALYKCEAVNKVGRGE 540  
Qy 541 RVLSFHVTRGPETITLOPDMQPTQESVSLMCTADRSTFENLTWYKLGQPPQLPIHVGELPT 600  
Db 541 RVLSFHVTRGPETITLOPDMQPTQESVSLMCTADRSTFENLTWYKLGQPPQLPIHVGELPT 600  
Qy 601 PCKNLDLTLWKLNATWFSNSTNDILIMELKNASLQDQGYVCLAQDRKTKRKHCVVRQLT 660  
Db 601 PCKNLDLTLWKLNATWFSNSTNDILIMELKNASLQDQGYVCLAQDRKTKRKHCVVRQLT 660  
Qy 661 VLERVAPITITGNLENQTTISGESIEVSCASGNPPQIMWPKONETLVEDSGIVLKDGNR 720  
Db 661 VLERVAPITITGNLENQTTISGESIEVSCASGNPPQIMWPKONETLVEDSGIVLKDGNR 720  
Qy 720 NLTIIRVRKDEGLYTCQACSVLGCACVAFVIEGAEQKTNLEIILVGTAVIAMPFWL 780  
Db 720 NLTIIRVRKDEGLYTCQACSVLGCACVAFVIEGAEQKTNLEIILVGTAVIAMPFWL 780  
Qy 781 LLVITILTVKVRANGELKTYLSIVMDPDELDELDEHCERLPYDASKWEPDRDLKGLKPL 840  
Db 781 LLVITILTVKVRANGELKTYLSIVMDPDELDELDEHCERLPYDASKWEPDRDLKGLKPL 840  
Qy 841 GRGAFQGVIEADAFGIDKATCTRTAVVVKMKEGATHSEHRALMSELKILIHGHILNVN 900  
Db 841 GRGAFQGVIEADAFGIDKATCTRTAVVVKMKEGATHSEHRALMSELKILIHGHILNVN 900  
Qy 901 LLGACTKPGGLMVI VEFCKFGNLSYLRSKRNEFVVKTKGARFRQGDYVGAIPVDLK 960  
Db 901 LLGACTKPGGLMVI VEFCKFGNLSYLRSKRNEFVVKTKGARFRQGDYVGAIPVDLK 960  
Qy 961 RRLDSITSSQSSASSGFGVEEKSLSDVVEEBEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
Db 961 RRLDSITSSQSSASSGFGVEEKSLSDVVEEBEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
Qy 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPKWMAPETIFDR 1080  
Db 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPKWMAPETIFDR 1080  
Qy 1081 VYTIQSDWMSFGVLLWEIFSLGASPGVKIDEEFCRRLEKGTMRAPDYTTPEMYOTML 1140  
Db 1081 VYTIQSDWMSFGVLLWEIFSLGASPGVKIDEEFCRRLEKGTMRAPDYTTPEMYOTML 1140  
Qy 1141 DCWHGEPSPQPTSELSVEHLGNLLQANAQDGKDYIVLPISSETLSMEEDSGLSLPTSPVS 1200  
Db 1141 DCWHGEPSPQPTSELSVEHLGNLLQANAQDGKDYIVLPISSETLSMEEDSGLSLPTSPVS 1200  
Qy 1201 CMEEEVCDPKFHYDNTAGISQYLQNSKRKRSPVSVKTFEDIPLIEEPVKVIPPDDNOTDS 1260  
Db 1201 CMEEEVCDPKFHYDNTAGISQYLQNSKRKRSPVSVKTFEDIPLIEEPVKVIPPDDNOTDS 1260  
Qy 1261 GWVLASEELKLTLEDRTKLSPSFGGMPVSKSRVASESGNSQTSYQSGYHSDDDTTTVYS 1320  
Db 1261 GWVLASEELKLTLEDRTKLSPSFGGMPVSKSRVASESGNSQTSYQSGYHSDDDTTTVYS 1320  
Qy 1321 SEAEALLKLTLEIGVQTGSTAQILQPDGTTLSPPV 1356  
Db 1321 SEAEALLKLTLEIGVQTGSTAQILQPDGTTLSPPV 1356

RESULT 11  
US-10-741-600-1471  
; Sequence 1471, Application US/10741600  
; Publication No. US20050026169A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001499  
; CURRENT APPLICATION NUMBER: US/10/741,600  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 73997  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 1471  
; LENGTH: 1356  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-741-600-1471

Query Match 99.9%; Score 7092; DB 5; Length 1356;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESKVLAAVLCVETRAASVGLPSVSLDPLRLSIQKDLITIKANTTLQITCRGQDLD 60  
Db 1 MQSKVLLAAVLCVETRAASVGLPSVSLDPLRLSIQKDLITIKANTTLQITCRGQDLD 60

Qy 61 WLPNNSGSEQRVEVTECSGDLFCKTLTTPKVI GNDTGAYKCFYRETDLASVIYVVD 120  
Db 61 WLPNNSGSEQRVEVTECSGDLFCKTLTTPKVI GNDTGAYKCFYRETDLASVIYVVD 120

Qy 121 YRPFPTASVDQGVVITENKNTVVI PCLGSI SNLNVSLCARYPEKRVDPGNRISWD 180  
Db 121 YRPFPTASVDQGVVITENKNTVVI PCLGSI SNLNVSLCARYPEKRVDPGNRISWD 180

Qy 181 SKKGFTIPSYMI SYAGMVFCEAKINDESQSI MYIVVVVGYRIYDVVLSPSHGIELSVGE 240  
Db 181 SKKGFTIPSYMI SYAGMVFCEAKINDESQSI MYIVVVVGYRIYDVVLSPSHGIELSVGE 240

Qy 241 KVLNCTARTELNVGIDFNWEYPSKHOKKLVNRDLKTQSGEMKKFLSTLIDGVTRS 300  
Db 241 KVLNCTARTELNVGIDFNWEYPSKHOKKLVNRDLKTQSGEMKKFLSTLIDGVTRS 300

Qy 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIPAKYLVGPPP 360  
Db 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIPAKYLVGPPP 360

Qy 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTGNTVILTNPI SKESQSHVSVLVVYVP 420  
Db 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTGNTVILTNPI SKESQSHVSVLVVYVP 420

Qy 421 PQIGKSLIPVDSYQYGTQTTLCTVYAI PPHPHIHWYQLEBECANBPQAVSVTNPY 480  
Db 421 PQIGKSLIPVDSYQYGTQTTLCTVYAI PPHPHIHWYQLEBECANBPQAVSVTNPY 480

Qy 481 PCEWMSVEDFOGKNLEVNKNQFALIEGKNKVTSTLVQAANVSALYKCEAVNKVGRGE 540  
Db 481 PCEWMSVEDFOGKNLEVNKNQFALIEGKNKVTSTLVQAANVSALYKCEAVNKVGRGE 540

Qy 541 RVISFHVTRGPEITLPDMQPTQESVSLWCTADRSTFENLTWYKGPQLPHVGELEPT 600  
Db 541 RVISFHVTRGPEITLPDMQPTQESVSLWCTADRSTFENLTWYKGPQLPHVGELEPT 600

Qy 601 PUCKNLDLTKLNATWFSNSTDILIMELKNASLDQGDYVCLAQRKTKGRHCVVRQLT 660  
Db 601 PUCKNLDLTKLNATWFSNSTDILIMELKNASLDQGDYVCLAQRKTKGRHCVVRQLT 660

Qy 661 VLBERVPTITGNLENQTTISGESI EYVSTASGNPPQIMWFKDNETLVEDSGVLKDGNR 720  
Db 661 VLBERVPTITGNLENQTTISGESI EYVSTASGNPPQIMWFKDNETLVEDSGVLKDGNR 720

Qy 721 NLTI RVRKEDEGLYTCQACSVLGCARVEAFFIIEGAQKTNLEIIILVGTAVIAMFFWL 780

Db 721 NLTI RVRKEDEGLYTCQACSVLGCARVEAFFIIEGAQKTNLEIIILVGTAVIAMFFWL 780  
Qy 781 LVVILITVRKANGGELKTCYLSIVMDPDELPLDEHCERLPYDASKWEPRDRKLKGLKPL 840  
Db 781 LVVILITVRKANGGELKTCYLSIVMDPDELPLDEHCERLPYDASKWEPRDRKLKGLKPL 840  
Qy 841 GRGAFQGVIEADAFGIDKTATCTAVAKMLKEGATHSEHRALMSELKILIHIGHLNVVN 900  
Db 841 GRGAFQGVIEADAFGIDKTATCTAVAKMLKEGATHSEHRALMSELKILIHIGHLNVVN 900  
Qy 901 LIGACTKPGGPLMVI VEFCKPFGNLSYLSRKNRNEFVYKTKGARFRQGDYVGAIPVDLK 960  
Db 901 LIGACTKPGGPLMVI VEFCKPFGNLSYLSRKNRNEFVYKTKGARFRQGDYVGAIPVDLK 960  
Qy 961 RRLDSITSSQSSASSGCFVEEKSLSVDEEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
Db 961 RRLDSITSSQSSASSGCFVEEKSLSVDEEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
Qy 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDVVRKGDARLPLKWWAPETIFDR 1080  
Db 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDVVRKGDARLPLKWWAPETIFDR 1080  
Qy 1081 VYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRRLKEGTRMRAPDYTTPEMYQTML 1140  
Db 1081 VYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRRLKEGTRMRAPDYTTPEMYQTML 1140  
Qy 1141 DCHWGEPSPORPFSSELVEHLGNLLQANAQODGKYIVLPISETLSNEEDSGLSLPTSPVS 1200  
Db 1141 DCHWGEPSPORPFSSELVEHLGNLLQANAQODGKYIVLPISETLSNEEDSGLSLPTSPVS 1200  
Qy 1201 CMEEEBVCDPKPHYDNTAGISQYVLSNKRKSPVSVKTFEDIPLEPEVKVIPDDNQDTS 1260  
Db 1201 CMEEEBVCDPKPHYDNTAGISQYVLSNKRKSPVSVKTFEDIPLEPEVKVIPDDNQDTS 1260  
Qy 1261 GMVLAEEELKTLIEDRTKLSFGMVPVPSKRSRESVASEGNSQTSYGYSQSHSDDTDTTVYS 1320  
Db 1261 GMVLAEEELKTLIEDRTKLSFGMVPVPSKRSRESVASEGNSQTSYGYSQSHSDDTDTTVYS 1320  
Qy 1321 SBEAEELLKLEIGVQGTGTAQI LQPDGTTLSPPV 1356  
Db 1321 SBEAEELLKLEIGVQGTGTAQI LQPDGTTLSPPV 1356

RESULT 12  
US-10-926-806-10  
; Sequence 10, Application US/10926806  
; Publication No. US20050096257A1  
; GENERAL INFORMATION:  
; APPLICANT: SHIMA, DAVID  
; APPLICANT: CALIAS, PERRY  
; APPLICANT: ADAMS, ANTHONY P.  
; TITLE OF INVENTION: COMBINATION THERAPY FOR THE TREATMENT OF OCULAR  
; TITLE OF INVENTION: NEOVASCULAR DISORDERS  
; FILE REFERENCE: 112089.184 EYE-013  
; CURRENT APPLICATION NUMBER: US/10/926,806  
; CURRENT FILING DATE: 2004-08-26  
; PRIOR APPLICATION NUMBER: 60/556,837  
; PRIOR FILING DATE: 2004-03-26  
; PRIOR APPLICATION NUMBER: 60/498,407  
; PRIOR FILING DATE: 2003-08-27  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 10  
; LENGTH: 1356  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-926-806-10

Query Match 99.9%; Score 7092; DB 5; Length 1356;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESKVLAVLWLCVETRAASVGLPSVSLDLPRLSIQKIDILTIKANTTLOITCRGQDLD 60  
DB 1 MQSKVLLAVLWLCVETRAASVGLPSVSLDLPRLSIQKIDILTIKANTTLOITCRGQDLD 60  
QY 61 WLPNNGSGSQREVEVTECSGDLFCCKTLTIIPKVIIGNDTGAYKCFYRETDLASVIYVYVD 120  
DB 61 WLPNNGSGSQREVEVTECSGDLFCCKTLTIIPKVIIGNDTGAYKCFYRETDLASVIYVYVD 120  
QY 121 YRSPFIASVSDOHGVVYITENKNTKNTVPCIGSISNLNLSICARYPKRFPVDPGNRISWD 180  
DB 121 YRSPFIASVSDOHGVVYITENKNTKNTVPCIGSISNLNLSICARYPKRFPVDPGNRISWD 180  
QY 181 SKKGFTIPSYMIYAGMVFCEAKINDESYQIMIVVVGRIYDVVLSPSHGIELSVE 240  
DB 181 SKKGFTIPSYMIYAGMVFCEAKINDESYQIMIVVVGRIYDVVLSPSHGIELSVE 240  
QY 241 KLVLNCTARTELANVGIDFNWEPSSKHQKLVNRDLKTQSGSEMKKFLSTLTIDGVTRS 300  
DB 241 KLVLNCTARTELANVGIDFNWEPSSKHQKLVNRDLKTQSGSEMKKFLSTLTIDGVTRS 300  
QY 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIPAKYLGYP 360  
DB 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIPAKYLGYP 360  
QY 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTGNTVILTNPISKEQSHVSLVYVP 420  
DB 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTGNTVILTNPISKEQSHVSLVYVP 420  
QY 421 PQIGESLISVDSYQYGTOTLACTVVAIPPPHHIHWMOLEECANESQASVTNYP 480  
DB 421 PQIGESLISVDSYQYGTOTLACTVVAIPPPHHIHWMOLEECANESQASVTNYP 480  
QY 481 PCEWRSVEDPQGNKIEVNKNQFALIEGKNTVSTLVIQAAVNSALYKCEAVNKVGRGE 540  
DB 481 PCEWRSVEDPQGNKIEVNKNQFALIEGKNTVSTLVIQAAVNSALYKCEAVNKVGRGE 540  
QY 541 RVISFHVTRGPEITLPQDMQTEQESVSLWCTADRTFENLTWYKLGPOPLPIHVGELPT 600  
DB 541 RVISFHVTRGPEITLPQDMQTEQESVSLWCTADRTFENLTWYKLGPOPLPIHVGELPT 600  
QY 601 PVCKNLOTLMKLNATMFSNSTNDILIMELKNASIQDQDYVCLAQDRKTKRHCVRQLT 660  
DB 601 PVCKNLOTLMKLNATMFSNSTNDILIMELKNASIQDQDYVCLAQDRKTKRHCVRQLT 660  
QY 661 VLERVAPTITGNLENQTTIGESIEVSTAGSNPPQIMWPKNETLVEDSGIVLKDGNR 720  
DB 661 VLERVAPTITGNLENQTTIGESIEVSTAGSNPPQIMWPKNETLVEDSGIVLKDGNR 720  
QY 721 NLTIIRVRKEDGLYTQACSVLCAKVEAFPIIEGAQEKTNLEIILVGTAVTAMFPL 780  
DB 721 NLTIIRVRKEDGLYTQACSVLCAKVEAFPIIEGAQEKTNLEIILVGTAVTAMFPL 780  
QY 781 LLVLIILTVKANGELKTGYLSIVMDPDELPLDEHCEPLPYDASKWEPDRDLKLGKPL 840  
DB 781 LLVLIILTVKANGELKTGYLSIVMDPDELPLDEHCEPLPYDASKWEPDRDLKLGKPL 840  
QY 841 GRGAGQVIEADAFGIDKATCRTVAVKMLKEGATHSEHRAIMSELKILIHGHILNVN 900  
DB 841 GRGAGQVIEADAFGIDKATCRTVAVKMLKEGATHSEHRAIMSELKILIHGHILNVN 900  
QY 901 LLGACTPFGGLMWIVBECKFNSTYLRSKRNEFVYKTKGARFQCKDYVGAIPVDLK 960  
DB 901 LLGACTPFGGLMWIVBECKFNSTYLRSKRNEFVYKTKGARFQCKDYVGAIPVDLK 960  
QY 961 RRLDSITSSQSSASSGFVEEKSLSVDEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
DB 961 RRLDSITSSQSSASSGFVEEKSLSVDEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020  
QY 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR 1080  
DB 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR 1080  
QY 1081 VYTIQSDWVSGVLLWEIFSLGASPYGVKIDEBFCRLKEGTRMRAPDYTPPMYQTML 1140

DB 1081 VYTIQSDWVSGVLLWEIFSLGASPYGVKIDEBFCRLKEGTRMRAPDYTPPMYQTML 1140  
QY 1141 DCWHGEPSPQRTTSELVEHLGNLLQANAQQDGKDYIVLPISSETLSMBEDSGLSLPTSPVS 1200  
DB 1141 DCWHGEPSPQRTTSELVEHLGNLLQANAQQDGKDYIVLPISSETLSMBEDSGLSLPTSPVS 1200  
QY 1201 CMEEERVCDFKHFYDNTAGISQYLONSKRKRSPVSKTFFEDIPLEEBEVKVIPODNOTDS 1260  
DB 1201 CMEEERVCDFKHFYDNTAGISQYLONSKRKRSPVSKTFFEDIPLEEBEVKVIPODNOTDS 1260  
QY 1261 GWVLASBELKLTEDRTKLSPSFGMVPKSKRESVASEGNSQTSYQSGYHSDDDTDTTVYS 1320  
DB 1261 GWVLASBELKLTEDRTKLSPSFGMVPKSKRESVASEGNSQTSYQSGYHSDDDTDTTVYS 1320  
QY 1321 SEEAELLKLIIEIGVQTGSTAQILQPDGTTLSPPV 1356  
DB 1321 SEEAELLKLIIEIGVQTGSTAQILQPDGTTLSPPV 1356  
RESULT 13  
US-10-824-982-2  
; Sequence 2, Application US/10824982  
; Publication No. US2005019742A1  
; GENERAL INFORMATION:  
; APPLICANT: Inventors: Michele A. McTigue, Steven L. Bender, Allen Borchardt, Robert  
; APPLICANT: S.  
; APPLICANT: Kania, Chris Pinko, John A. Wickersham  
; TITLE OF INVENTION: Crystal Structure of VEGFRK2: Ligand Complexes and  
; FILE OF INVENTION: Methods of Use Thereof  
; FILE REFERENCE: PC19173A  
; CURRENT APPLICATION NUMBER: US/10/824,982  
; CURRENT FILING DATE: 2004-04-15  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 1356  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-824-982-2  
Query Match 99.9%; Score 7092; DB 5; Length 1356;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MESKVLAVLWLCVETRAASVGLPSVSLDLPRLSIQKIDILTIKANTTLOITCRGQDLD 60  
DB 1 MQSKVLLAVLWLCVETRAASVGLPSVSLDLPRLSIQKIDILTIKANTTLOITCRGQDLD 60  
QY 61 WLPNNGSGSQREVEVTECSGDLFCCKTLTIIPKVIIGNDTGAYKCFYRETDLASVIYVYVD 120  
DB 61 WLPNNGSGSQREVEVTECSGDLFCCKTLTIIPKVIIGNDTGAYKCFYRETDLASVIYVYVD 120  
QY 121 YRSPFIASVSDOHGVVYITENKNTKNTVPCIGSISNLNLSICARYPKRFPVDPGNRISWD 180  
DB 121 YRSPFIASVSDOHGVVYITENKNTKNTVPCIGSISNLNLSICARYPKRFPVDPGNRISWD 180  
QY 181 SKKGFTIPSYMIYAGMVFCEAKINDESYQIMIVVVGRIYDVVLSPSHGIELSVE 240  
DB 181 SKKGFTIPSYMIYAGMVFCEAKINDESYQIMIVVVGRIYDVVLSPSHGIELSVE 240  
QY 241 KLVLNCTARTELANVGIDFNWEPSSKHQKLVNRDLKTQSGSEMKKFLSTLTIDGVTRS 300  
DB 241 KLVLNCTARTELANVGIDFNWEPSSKHQKLVNRDLKTQSGSEMKKFLSTLTIDGVTRS 300  
QY 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIPAKYLGYP 360  
DB 301 DOGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIPAKYLGYP 360  
QY 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTGNTVILTNPISKEQSHVSLVYVP 420  
DB 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSRDGTGNTVILTNPISKEQSHVSLVYVP 420

```
Qy 421 POIGKSLSPVDSYQYGTQTTLCTVYAI PPPHHIHWYQLEBECANEPFSAVSVTNPY 480
Db 421 POIGKSLSPVDSYQYGTQTTLCTVYAI PPPHHIHWYQLEBECANEPFSAVSVTNPY 480
Qy 481 PCEWRSVDFOGNKEVKNQFALLIEGKNTVSTLVIQAANVSALYKCEAVNKVGRGE 540
Db 481 PCEWRSVDFOGNKEVKNQFALLIEGKNTVSTLVIQAANVSALYKCEAVNKVGRGE 540
Qy 541 RVISFHVTRGPEITLQPDMPTEOSVSLCTADRSTFENLTWYKLGQPOLPIHVHGLPT 600
Db 541 RVISFHVTRGPEITLQPDMPTEOSVSLCTADRSTFENLTWYKLGQPOLPIHVHGLPT 600
Qy 601 PCKNLDLTKLKNATMFSNSTNDILIMELKNASLQDQGVYCLAQDRKTKRHCVRQLT 660
Db 601 PCKNLDLTKLKNATMFSNSTNDILIMELKNASLQDQGVYCLAQDRKTKRHCVRQLT 660
Qy 661 VLERVAPITICNLENQTTISGESIEVSTAGNPPQIMWFKONETLVEDSGIVLKDGNR 720
Db 661 VLERVAPITICNLENQTTISGESIEVSTAGNPPQIMWFKONETLVEDSGIVLKDGNR 720
Qy 721 NLITRRVRKDEGLYTQACSVLGCACVAFPIIEGAQEKTNLEIILVGTAVIAMPFWL 780
Db 721 NLITRRVRKDEGLYTQACSVLGCACVAFPIIEGAQEKTNLEIILVGTAVIAMPFWL 780
Qy 781 LLVILITVRKANGELKTYLSIVMDPDELPLDEHCERLPYDASKWFFPRDRLLKLGKPL 840
Db 781 LLVILITVRKANGELKTYLSIVMDPDELPLDEHCERLPYDASKWFFPRDRLLKLGKPL 840
Qy 841 GRGAFQVIEADAFIDKATCRTVAVKMLKEGATHSEHRALMSLKIILHGHILNVN 900
Db 841 GRGAFQVIEADAFIDKATCRTVAVKMLKEGATHSEHRALMSLKIILHGHILNVN 900
Qy 901 LLGACTKPGGLPMVIEFCKFNGMLSTYLRSKRNEFVPYKTKGARFRQGDYVGAIPVDLK 960
Db 901 LLGACTKPGGLPMVIEFCKFNGMLSTYLRSKRNEFVPYKTKGARFRQGDYVGAIPVDLK 960
Qy 961 RRLDSITSSQSSASSGFVEEKSLSDBEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020
Db 961 RRLDSITSSQSSASSGFVEEKSLSDBEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA 1020
Qy 1021 SRKCIHRDLAARNILISEKNVVKICDFGLARDIYKOPDYVRKGDARLPLKWMAPETIFDR 1080
Db 1021 SRKCIHRDLAARNILISEKNVVKICDFGLARDIYKOPDYVRKGDARLPLKWMAPETIFDR 1080
Qy 1081 VYTIQSDVMSFGVLLWEIIFSLGASPGVGVKIDEEFCRLKEGTRMAPDYTTPEMYQTM 1140
Db 1081 VYTIQSDVMSFGVLLWEIIFSLGASPGVGVKIDEEFCRLKEGTRMAPDYTTPEMYQTM 1140
Qy 1141 DCWHGSPQRPTFSELVEHLGNLLQANAQDQGDYIVLPISETLSMEEDSGLSLPTSPVS 1200
Db 1141 DCWHGSPQRPTFSELVEHLGNLLQANAQDQGDYIVLPISETLSMEEDSGLSLPTSPVS 1200
Qy 1201 CMEEBEEVCDPKFYDNTAGISQYLQNSKRKRSPVSVKTFPEDIPLIEPPEVKVIPDDNQDTS 1260
Db 1201 CMEEBEEVCDPKFYDNTAGISQYLQNSKRKRSPVSVKTFPEDIPLIEPPEVKVIPDDNQDTS 1260
Qy 1261 GMYLASELTKLDRTKLSPFGMVPKSRRESVASEGNSQTSYGSHSDDDTDTTVYS 1320
Db 1261 GMYLASELTKLDRTKLSPFGMVPKSRRESVASEGNSQTSYGSHSDDDTDTTVYS 1320
Qy 1321 SEEAELLKLEIGVQTGSTAQIILQPDSGTTLSPPPV 1356
Db 1321 SEEAELLKLEIGVQTGSTAQIILQPDSGTTLSPPPV 1356
```

RESULT 14

```
US-11-021-951-115
; Sequence 115, Application US/11021951
; Publication No. US20050175581A1
; GENERAL INFORMATION:
; APPLICANT: HAUPTS, Ulrich
; APPLICANT: KOLTERMANN, Andre
; APPLICANT: SCHEIDIG, Andreas
```

```
; APPLICANT: VOTSMEIER, Christian
; APPLICANT: Ketting, Ulrich
; APPLICANT: COCO, Wayne Michael
; TITLE OF INVENTION: New Biological Entities And The Pharmaceutical
; TITLE OF INVENTION: And Diagnostic Use Thereof
; FILE REFERENCE: 04156.0002U5
; CURRENT APPLICATION NUMBER: US/11/021,951
; CURRENT FILING DATE: 2004-12-22
; PRIOR APPLICATION NUMBER: 10/872,198
; PRIOR FILING DATE: 2004-06-18
; PRIOR APPLICATION NUMBER: 60/543,518
; PRIOR FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: 60/524,960
; PRIOR FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: EP 04003058
; PRIOR FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: EP 03025871
; PRIOR FILING DATE: 2003-11-11
; PRIOR APPLICATION NUMBER: EP 03025851
; PRIOR FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: EP 03013819
; PRIOR FILING DATE: 2003-06-18
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 115
; LENGTH: 1356
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-021-951-115
```

```
Query Match 99.9%; Score 7092; DB 6; Length 1356;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MESKVLVALVMLCVETRAASVGLPSVSLDLPKLSIQKDILITKANTTLQITCRGQDLD 60
Db 1 MOSKVLVALVMLCVETRAASVGLPSVSLDLPKLSIQKDILITKANTTLQITCRGQDLD 60
Qy 61 WLPNNQSGSEORVEVTECSGLFCCKTLTPKVIGNDTGAYKCFYRETDLASVIVYVQD 120
Db 61 WLPNNQSGSEORVEVTECSGLFCCKTLTPKVIGNDTGAYKCFYRETDLASVIVYVQD 120
Qy 121 YRSPFIASVSDQHGVMYITENKNTVVIICLGSISNLNLSLCARYPEKFPVDPGNRISWD 180
Db 121 YRSPFIASVSDQHGVMYITENKNTVVIICLGSISNLNLSLCARYPEKFPVDPGNRISWD 180
Qy 181 SKKGFTIPSYMISYAGWFCFAKINDESYQSIMYIVVVVGYRIYDVVLSPSHGIELSVGE 240
Db 181 SKKGFTIPSYMISYAGWFCFAKINDESYQSIMYIVVVVGYRIYDVVLSPSHGIELSVGE 240
Qy 241 KLVLNCTARTELNVGIDFNWEYPSSKHQHKLVNRDLKTQSGSEMCKFLSTLTIDGVTRS 300
Db 241 KLVLNCTARTELNVGIDFNWEYPSSKHQHKLVNRDLKTQSGSEMCKFLSTLTIDGVTRS 300
Qy 301 DQGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIPAKYLGYPPP 360
Db 301 DQGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIPAKYLGYPPP 360
Qy 361 EIKWYKNGIPLSNNHTIKAGHVLTIMEVSEDTGNTVILTNPISKEKQSHVVSVLVVYVP 420
Db 361 EIKWYKNGIPLSNNHTIKAGHVLTIMEVSEDTGNTVILTNPISKEKQSHVVSVLVVYVP 420
Qy 421 POIGKSLSPVDSYQYGTQTTLCTVYAI PPPHHIHWYQLEBECANEPFSAVSVTNPY 480
Db 421 POIGKSLSPVDSYQYGTQTTLCTVYAI PPPHHIHWYQLEBECANEPFSAVSVTNPY 480
Qy 481 PCEWRSVDFOGNKEVKNQFALLIEGKNTVSTLVIQAANVSALYKCEAVNKVGRGE 540
Db 481 PCEWRSVDFOGNKEVKNQFALLIEGKNTVSTLVIQAANVSALYKCEAVNKVGRGE 540
Qy 541 RVISFHVTRGPEITLQPDMPTEOSVSLCTADRSTFENLTWYKLGQPOLPIHVHGLPT 600
Db 541 RVISFHVTRGPEITLQPDMPTEOSVSLCTADRSTFENLTWYKLGQPOLPIHVHGLPT 600
```

QY 601 PVCKNLDLTLKLNATMPSNSTNDILIMELKNASLODQDGYVCLAQDRKTKKRGHCVRQLT 660  
DB 601 PVCKNLDLTLKLNATMPSNSTNDILIMELKNASLODQDGYVCLAQDRKTKKRGHCVRQLT 660  
QY 661 VLVRVAPITIGNLENQTTISGESIEVSTAGNPPQIMWFKDNETLVEDSGIVLKDGNR 720  
DB 661 VLVRVAPITIGNLENQTTISGESIEVSTAGNPPQIMWFKDNETLVEDSGIVLKDGNR 720  
QY 721 NLTIIRVRKEDEGLYTQACSVLGCACVAFIIEGAQKTNLEIILVGTAVIAMPFWL 780  
DB 721 NLTIIRVRKEDEGLYTQACSVLGCACVAFIIEGAQKTNLEIILVGTAVIAMPFWL 780  
QY 781 LLVILIRTVKANGELKTYLSIVMDPDELPHDECHERLPYDASKWFFPRDLKLGKPL 840  
DB 781 LLVILIRTVKANGELKTYLSIVMDPDELPHDECHERLPYDASKWFFPRDLKLGKPL 840  
QY 841 GRGAFQVIEADAFGIDKATCRIVAVKMLKEGATHSEHRALMSELKILIHGHILNVN 900  
DB 841 GRGAFQVIEADAFGIDKATCRIVAVKMLKEGATHSEHRALMSELKILIHGHILNVN 900  
QY 901 LLGACTKPGGLMVIIEFCKFNGNSTYLRSKRNEFVYKTKGARFRQKDVGAIPVDLK 960  
DB 901 LLGACTKPGGLMVIIEFCKFNGNSTYLRSKRNEFVYKTKGARFRQKDVGAIPVDLK 960  
QY 961 RRLDSITSSQSSASSGFVEEKSLSDBVEEBAPEDLYKDFLTLEHLICYSFOVAKGMEFLA 1020  
DB 961 RRLDSITSSQSSASSGFVEEKSLSDBVEEBAPEDLYKDFLTLEHLICYSFOVAKGMEFLA 1020  
QY 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPYVRKGDARLPKWMAPETIFDR 1080  
DB 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPYVRKGDARLPKWMAPETIFDR 1080  
QY 1081 VYTIQSDWMSFVLLWEIFSLGASPYGVKIDEEFCRLKEGTRMRAPDYTPPMYQTM 1140  
DB 1081 VYTIQSDWMSFVLLWEIFSLGASPYGVKIDEEFCRLKEGTRMRAPDYTPPMYQTM 1140  
QY 1141 DCMWGESQRTPELVEHLGNLLOANAQDGKDYIVLPISETLSMBEDSGLSLPTSPVS 1200  
DB 1141 DCMWGESQRTPELVEHLGNLLOANAQDGKDYIVLPISETLSMBEDSGLSLPTSPVS 1200  
QY 1201 CMEEEEECDPKFHYDNTAGISQYLONSKRSPVSVKTFEDIPLEEPVKVIPPDDNQDTS 1260  
DB 1201 CMEEEEECDPKFHYDNTAGISQYLONSKRSPVSVKTFEDIPLEEPVKVIPPDDNQDTS 1260  
QY 1261 GMVLASBELKTLERDRTKLSPSFGMVPSKSRVASEGNSQTSYGSGYHSDDDTDTTVYS 1320  
DB 1261 GMVLASBELKTLERDRTKLSPSFGMVPSKSRVASEGNSQTSYGSGYHSDDDTDTTVYS 1320  
QY 1321 SEAEALLKLIBIGVQTGSTAQILOPDSGTTLSPPV 1356  
DB 1321 SEAEALLKLIBIGVQTGSTAQILOPDSGTTLSPPV 1356

RESULT 15  
US-09-969-037-7  
; Sequence 7, Application US/09969037  
; Publication No. US2003002247A1  
; GENERAL INFORMATION:  
; APPLICANT: KYOKA HAKKO KOGYO CO., LTD.  
; TITLE OF INVENTION: Substance which inhibits binding of information transfer molecule  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/09/969,037  
; PRIOR FILING DATE: 2001-10-03  
; PRIOR APPLICATION NUMBER: JP 2000-303694  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: US 60/263,512  
; PRIOR FILING DATE: 2001-01-24  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 1356

; TYPE: PRT  
; ORGANISM: Human  
US-09-969-037-7

Query Match 99.9%; Score 7091; DB 3; Length 1356;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1354; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESKVLALVALWLCVETRAASVGLPSVLDLPLRLSIQDKILTIKANTTLQITTCRGQRDL 60  
DB 1 MQSKVLALVALWLCVETRAASVGLPSVLDLPLRLSIQDKILTIKANTTLQITTCRGQRDL 60  
QY 61 WLNNNGSGEORVEVTECDGLPCKTLTIPKVIIGNDTGAYKCFYRETDLASVIVVYQD 120  
DB 61 WLNNNGSGEORVEVTECDGLPCKTLTIPKVIIGNDTGAYKCFYRETDLASVIVVYQD 120  
QY 121 YRSPFIASVSDQHGVIYITENKNTVVIICLGSISNLNLSLCARYPEKRFVPDGNRISWD 180  
DB 121 YRSPFIASVSDQHGVIYITENKNTVVIICLGSISNLNLSLCARYPEKRFVPDGNRISWD 180  
QY 181 SKGFTIPSMISVAGMVFCEAKINDESYOSIMYIVVVVGYRIYDVVLSPSHGIELSVGE 240  
DB 181 SKGFTIPSMISVAGMVFCEAKINDESYOSIMYIVVVVGYRIYDVVLSPSHGIELSVGE 240  
QY 241 KVLNCTARTELNVGIDFNWEYPSKHOKKLVNRDLKTOSGSEMKKFLSTLTIDGVT 300  
DB 241 KVLNCTARTELNVGIDFNWEYPSKHOKKLVNRDLKTOSGSEMKKFLSTLTIDGVT 300  
QY 301 DOGLYTCAASGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIPAKLYGYPPP 360  
DB 301 DOGLYTCAASGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIPAKLYGYPPP 360  
QY 361 EIKWYKNGIPLESNHTIKAGHVLTIMESVSDTGNVTIILTNPISEKQSHVSLVYVP 420  
DB 361 EIKWYKNGIPLESNHTIKAGHVLTIMESVSDTGNVTIILTNPISEKQSHVSLVYVP 420  
QY 421 PQIGESKLSIPVDSYQGTQTLTCTVYAIIPPHHIIHWYQLEBECANESQAVSTN 480  
DB 421 PQIGESKLSIPVDSYQGTQTLTCTVYAIIPPHHIIHWYQLEBECANESQAVSTN 480  
QY 481 PCSEWRSVDFQGNKIEVKNQFALIEGKNTVSTLVIQANVSALYKCEAVNKVGRGE 540  
DB 481 PCSEWRSVDFQGNKIEVKNQFALIEGKNTVSTLVIQANVSALYKCEAVNKVGRGE 540  
QY 541 RVISFHVTRGPEITLQDPMQPTQESVSLMCTADRSTFENLTWYKLGQPLPIHVBELPT 600  
DB 541 RVISFHVTRGPEITLQDPMQPTQESVSLMCTADRSTFENLTWYKLGQPLPIHVBELPT 600  
QY 601 PVCKNLDLTLKLNATMPSNSTNDILIMELKNASLODQDGYVCLAQDRKTKKRGHCVRQLT 660  
DB 601 PVCKNLDLTLKLNATMPSNSTNDILIMELKNASLODQDGYVCLAQDRKTKKRGHCVRQLT 660  
QY 661 VLVRVAPITIGNLENQTTISGESIEVSTAGNPPQIMWFKDNETLVEDSGIVLKDGNR 720  
DB 661 VLVRVAPITIGNLENQTTISGESIEVSTAGNPPQIMWFKDNETLVEDSGIVLKDGNR 720  
QY 721 NLTIIRVRKEDEGLYTQACSVLGCACVAFIIEGAQKTNLEIILVGTAVIAMPFWL 780  
DB 721 NLTIIRVRKEDEGLYTQACSVLGCACVAFIIEGAQKTNLEIILVGTAVIAMPFWL 780  
QY 781 LLVILIRTVKANGELKTYLSIVMDPDELPHDECHERLPYDASKWFFPRDLKLGKPL 840  
DB 781 LLVILIRTVKANGELKTYLSIVMDPDELPHDECHERLPYDASKWFFPRDLKLGKPL 840  
QY 841 GRGAFQVIEADAFGIDKATCRIVAVKMLKEGATHSEHRALMSELKILIHGHILNVN 900  
DB 841 GRGAFQVIEADAFGIDKATCRIVAVKMLKEGATHSEHRALMSELKILIHGHILNVN 900  
QY 901 LLGACTKPGGLMVIIEFCKFNGNSTYLRSKRNEFVYKTKGARFRQKDVGAIPVDLK 960  
DB 901 LLGACTKPGGLMVIIEFCKFNGNSTYLRSKRNEFVYKTKGARFRQKDVGAIPVDLK 960  
QY 961 RRLDSITSSQSSASSGFVEEKSLSDBVEEBAPEDLYKDFLTLEHLICYSFOVAKGMEFLA 1020

Db	961	RRLDSTSSQSSASSGFVEEKSLSDVEEEAPEDLYKDFLTLEHLICYSFQVAKGMEFLA	1020
Qy	1021	SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR	1080
Db	1021	SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR	1080
Qy	1081	VYTIQSDVWSFGVLLWEIFSLGASPYFGVKIDEEFCRRLKEGTRMRAPDYTTPEMYQTML	1140
Db	1081	VYTIQSDVWSFGVLLWEIFSLGASPYFGVKIDEEFCRRLKEGTRMRAPDYTTPEMYQTML	1140
Qy	1141	DCWHGEPSPQPTFSELVEHLGNLLQANAQQDGKDYIVLPISETLSMEEDSGLSLPTSPVS	1200
Db	1141	DCWHGEPSPQPTFSELVEHLGNLLQANAQQDGKDYIVLPISETLSMEEDSGLSLPTSPVS	1200
Qy	1201	CMEEEVCDPKFHYDNTAGISQYLQNSKRKSRPVSVKTFEDIPLPEEPEVKVIPDDNQDTS	1260
Db	1201	CMEEEVCDPKFHYDNTAGISQYLQNSKRKSRPVSVKTFEDIPLPEEPEVKVIPDDNQDTS	1260
Qy	1261	GMVLASEELKTLEDRTKLSPSFGMWPKSRESVASEGSNQTSQYQSGYHSDDDTDTTVYS	1320
Db	1261	GMVLASEELKTLEDRTKLSPSFGMWPKSRESVASEGSNQTSQYQSGYHSDDDTDTTVYS	1320
Qy	1321	SEEAELKLIEIGVQTGSTAQILQPDSGTTLSSPPV	1356
Db	1321	SEEAELKLIEIGVQTGSTAQILQPDSGTTLSSPPV	1356

Search completed: March 10, 2006, 19:06:49  
Job time : 139.671 secs



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2006, 19:00:41 ; Search time 14.4695 Seconds  
(without alignments)  
2608.526 Million cell updates/sec

Title: US-10-633-742-6  
Perfect score: 7095  
Sequence: 1 MESKVLAVLWLCVETRAA.....GSTAQILQPDGTTLSPPV 1356

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 161667 seqs, 27834885 residues

Total number of hits satisfying chosen parameters: 161667

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA New.\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pdb.\*  
2: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pdb.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pdb.\*  
4: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pdb.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pdb.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pdb.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pdb.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pdb.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7092.	99.9	1356	6	US-10-995-561-904
2	7092	99.9	1356	6	US-10-995-561-906
3	6806	95.9	1306	6	US-10-995-561-905
4	4020	56.7	764	7	US-11-075-047A-4
5	4017	56.6	767	7	US-11-043-693-2
6	4012	56.5	764	7	US-11-104-110-8
7	4012	56.5	764	7	US-11-104-111-28
8	3075	43.3	592	7	US-11-104-110-9
9	3075	43.3	592	7	US-11-104-111-29
10	2797.5	39.4	1368	7	US-11-043-693-34
11	2759	38.9	1363	7	US-11-076-427A-10
12	2759	38.9	1363	7	US-11-076-427A-121
13	2751.5	38.8	1298	7	US-11-076-427A-12
14	2751.5	38.8	1298	7	US-11-075-047A-6
15	2748.5	38.7	1338	7	US-11-109-156-23
16	2748.5	38.7	1338	7	US-11-075-047A-2
17	2746.5	38.7	1338	6	US-10-821-234-1622
18	2743.5	38.7	1362	7	US-11-043-693-33
19	2732	38.5	1363	7	US-11-043-693-32
20	1651	23.3	570	7	US-11-075-047A-8
21	1647	23.2	316	7	US-11-092-168-10
22	1315	18.5	583	7	US-11-075-047A-125
23	1315	18.5	583	7	US-11-075-047A-128
24	1255.5	17.7	777	7	US-11-043-693-3
25	1219.5	17.2	1451	6	US-10-995-561-829

26 1218.5 17.2 1089 7 US-11-148-770-32 Sequence 32, Appl  
27 1218.5 17.2 1089 7 US-11-076-427A-14 Sequence 14, Appl  
28 1218.5 17.2 1089 7 US-11-075-047A-117 Sequence 117, Appl  
29 1218.5 17.2 1089 7 US-11-177-894-26 Sequence 26, Appl  
30 1218.5 17.2 1089 7 US-11-177-894-27 Sequence 27, Appl  
31 1218.5 17.2 1089 7 US-11-177-894-28 Sequence 28, Appl  
32 1218.5 17.2 1089 7 US-11-177-894-29 Sequence 29, Appl  
33 1167 16.4 432 7 US-11-075-047A-16 Sequence 16, Appl  
34 1142 16.1 474 7 US-11-075-047A-20 Sequence 20, Appl  
35 1120.5 15.8 483 7 US-11-075-047A-14 Sequence 14, Appl  
36 1092.5 15.4 758 7 US-11-043-693-1 Sequence 1, Appl  
37 1088.5 15.3 477 7 US-11-075-047A-12 Sequence 12, Appl  
38 1057.5 14.9 1106 7 US-11-076-427A-16 Sequence 16, Appl  
39 1052.5 14.8 471 7 US-11-075-047A-119 Sequence 119, Appl  
40 1052.5 14.8 972 6 US-10-821-234-1587 Sequence 10, Appl  
41 1052 14.8 972 6 US-10-821-234-1587 Sequence 1587, Ap  
42 1034.5 14.6 976 7 US-11-148-770-31 Sequence 31, Appl  
43 1034.5 14.6 976 7 US-11-177-894-15 Sequence 15, Appl  
44 1034.5 14.6 976 7 US-11-177-894-16 Sequence 16, Appl  
45 1034.5 14.6 976 7 US-11-177-894-18 Sequence 18, Appl

## ALIGNMENTS

## RESULT 1

US-10-995-561-904  
; Sequence 904, Application US/10995561  
; Publication No. US2005027054A1  
; GENERAL INFORMATION: Michele et al.  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
; TITLE OF INVENTION: DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001559  
; CURRENT APPLICATION NUMBER: US/10/995,561  
; CURRENT FILING DATE: 2004-11-24  
; NUMBER OF SEQ ID NOS: 85702  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 904  
; LENGTH: 1356  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-995-561-904

Query Match 99.9%; Score 7092; DB 6; Length 1356;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESKVLAVLWLCVETRAASVGLPSVLDLPRLSIQKDLTIKANTTLOITCRGQDLD 60  
Db 1 MQSKVLAVLWLCVETRAASVGLPSVLDLPRLSIQKDLTIKANTTLOITCRGQDLD 60  
Qy 61 WLFPNNSQSEQRVEVTECSGLFCKTLTPKVGNDTGAYKCFYRETDLASVIYVVD 120  
Db 61 WLFPNNSQSEQRVEVTECSGLFCKTLTPKVGNDTGAYKCFYRETDLASVIYVVD 120  
Qy 121 YRSPFTASVSDQGVVYITENKNTVVPCLGSIINLSLCARYPKRFPVPGNRI SWD 180  
Db 121 YRSPFTASVSDQGVVYITENKNTVVPCLGSIINLSLCARYPKRFPVPGNRI SWD 180  
Qy 181 SKKGFTIPSMISYAGMVFCEAKINDESQYIMYVVVVGRIYDVVLSPSHGIELSVGE 240  
Db 181 SKKGFTIPSMISYAGMVFCEAKINDESQYIMYVVVVGRIYDVVLSPSHGIELSVGE 240  
Qy 241 KLVLNCTARTELNVGIDFNWYESSKHQKLVNRLKTQSGSEMKKFLSTLTIDGVTRS 300  
Db 241 KLVLNCTARTELNVGIDFNWYESSKHQKLVNRLKTQSGSEMKKFLSTLTIDGVTRS 300  
Qy 301 DOGLYTCASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVIRIPAKYLGYP 360  
Db 301 DOGLYTCASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVIRIPAKYLGYP 360

Qy 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSDTGNVTYVILTNPISEKQSHVSVLVVYP 420  
Db |||||  
Qy 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSDTGNVTYVILTNPISEKQSHVSVLVVYP 420  
Db |||||  
Qy 421 PQIGKSLISPVDSYQYGTQTTLCTVYAI PPPHHIHWYQLEBECANESQAVSVTNPY 480  
Db |||||  
Qy 421 PQIGKSLISPVDSYQYGTQTTLCTVYAI PPPHHIHWYQLEBECANESQAVSVTNPY 480  
Db |||||  
Qy 481 PCEWRSVDFQGNKIEVNKNQFALIEGKNKTSTLVIQANVSALYKCEAVNKVGRGE 540  
Db |||||  
Qy 481 PCEWRSVDFQGNKIEVNKNQFALIEGKNKTSTLVIQANVSALYKCEAVNKVGRGE 540  
Db |||||  
Qy 541 RVISFHVTRGPEITLQDMQPTQESVSLMCTADRSFTENLTWYKLGQPLPIHVGLPT 600  
Db |||||  
Qy 541 RVISFHVTRGPEITLQDMQPTQESVSLMCTADRSFTENLTWYKLGQPLPIHVGLPT 600  
Db |||||  
Qy 601 PCKNLDLTKLNATMFSNSTNDILIMELKNASLDQGDYVCLADRKTKKRHCVRQLT 660  
Db |||||  
Qy 601 PCKNLDLTKLNATMFSNSTNDILIMELKNASLDQGDYVCLADRKTKKRHCVRQLT 660  
Db |||||  
Qy 661 VLERVAPITIGNLENQTTSGIESIEVSCASGNPPPOIMWFKDNETLVEDSGIVLKDGNR 720  
Db |||||  
Qy 661 VLERVAPITIGNLENQTTSGIESIEVSCASGNPPPOIMWFKDNETLVEDSGIVLKDGNR 720  
Db |||||  
Qy 721 NLTIIRVRKDEGLYTCQACSVLGCAKVEAFFIIEGAQKTNLEIILVGTAVIAMFFWL 780  
Db |||||  
Qy 721 NLTIIRVRKDEGLYTCQACSVLGCAKVEAFFIIEGAQKTNLEIILVGTAVIAMFFWL 780  
Db |||||  
Qy 781 LLVILITVRKANGGELTKGYLSIVMDPDELPIDEHCHERLPYDASKWEPFRDLKLGKPL 840  
Db |||||  
Qy 781 LLVILITVRKANGGELTKGYLSIVMDPDELPIDEHCHERLPYDASKWEPFRDLKLGKPL 840  
Db |||||  
Qy 841 GRGAFQGVIBADAFIGDKTATCRVAVKMLKEGATHSEHRLMSELKILIHIGHLNVN 900  
Db |||||  
Qy 841 GRGAFQGVIBADAFIGDKTATCRVAVKMLKEGATHSEHRLMSELKILIHIGHLNVN 900  
Db |||||  
Qy 901 LLGACTKPGGLMVI VEFCKFGLNSTLYRSKRNEFPYKTKGARFRQGDYVGAIPVDLK 960  
Db |||||  
Qy 901 LLGACTKPGGLMVI VEFCKFGLNSTLYRSKRNEFPYKTKGARFRQGDYVGAIPVDLK 960  
Db |||||  
Qy 961 RRLDSITSSQSSASSGFVEBSKLSVDEEBEAPEDLYKDFLTLEHLICYSQVAKGMEFLA 1020  
Db |||||  
Qy 961 RRLDSITSSQSSASSGFVEBSKLSVDEEBEAPEDLYKDFLTLEHLICYSQVAKGMEFLA 1020  
Db |||||  
Qy 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPKAWAPETIFDR 1080  
Db |||||  
Qy 1021 SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPKAWAPETIFDR 1080  
Db |||||  
Qy 1081 VYTIQSDVWFGVLLWEIFSLGASPYGVKIDBEFCRRLEKGRMRAPDYTTPEMYQTM 1140  
Db |||||  
Qy 1081 VYTIQSDVWFGVLLWEIFSLGASPYGVKIDBEFCRRLEKGRMRAPDYTTPEMYQTM 1140  
Db |||||  
Qy 1141 DCWHGEPSPRTSELVEHIGNLQANQDGKDYIVLPISSETLSMEEDSGLSPTSPVS 1200  
Db |||||  
Qy 1141 DCWHGEPSPRTSELVEHIGNLQANQDGKDYIVLPISSETLSMEEDSGLSPTSPVS 1200  
Db |||||  
Qy 1201 CMEEEEECDPKFVNDTAGISQYLONSKRKSRPVSVKTFREDIPLERPEVAVIPDDNQDTS 1260  
Db |||||  
Qy 1201 CMEEEEECDPKFVNDTAGISQYLONSKRKSRPVSVKTFREDIPLERPEVAVIPDDNQDTS 1260  
Db |||||  
Qy 1261 GMVLASELKTLEDRTKLSPSFGGWPSKRSRESVASEGNSQTSQYSGYHSDDTDTTVYS 1320  
Db |||||  
Qy 1261 GMVLASELKTLEDRTKLSPSFGGWPSKRSRESVASEGNSQTSQYSGYHSDDTDTTVYS 1320  
Db |||||  
Qy 1321 SEBAELLKLEIGVQTGSTAQIQLQPSGTTLSPPV 1356  
Db |||||  
Qy 1321 SEBAELLKLEIGVQTGSTAQIQLQPSGTTLSPPV 1356  
Db |||||

RESULT 2

US-10-995-561-906

; Sequence 906, Application US/10995561

; Publication No. US20050272054A1

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF

; TITLE OF INVENTION: DETECTION AND USES THEREOF

; FILE REFERENCE: CU001559

; CURRENT APPLICATION NUMBER: US/10/995,561

; CURRENT FILING DATE: 2004-11-24

; NUMBER OF SEQ ID NOS: 85702

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 906

; LENGTH: 1356

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-995-561-906

Query Match 99.9%; Score 7092; DB 6; Length 1356;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESKVLALVALMVCVETRAASVGLPSVSLDLPRLSIQKDLITIKANTTILQITCRGORDLD 60

Db |||||

Qy 1 MQSKVLALVALMVCVETRAASVGLPSVSLDLPRLSIQKDLITIKANTTILQITCRGORDLD 60

Db |||||

Qy 61 WLPNNQSGSEQREVEVTECSGDLFCFKTLTI PKVIGNDTGAYKCFYRETDLASVIVVYQD 120

Db |||||

Qy 61 WLPNNQSGSEQREVEVTECSGDLFCFKTLTI PKVIGNDTGAYKCFYRETDLASVIVVYQD 120

Db |||||

Qy 121 YRSPFTASVSDQHGVIYITENKNTVVI PCLGSI SNLNSLCARYPEKRPVDPGNRISWD 180

Db |||||

Qy 121 YRSPFTASVSDQHGVIYITENKNTVVI PCLGSI SNLNSLCARYPEKRPVDPGNRISWD 180

Db |||||

Qy 181 SKGFTIPSMISYAGWFCFAKINDESQSIWV VVVVGRIYDVVLSPSHGIELSVGE 240

Db |||||

Qy 181 SKGFTIPSMISYAGWFCFAKINDESQSIWV VVVVGRIYDVVLSPSHGIELSVGE 240

Db |||||

Qy 241 KLVLNCTARTELNVGIDFNWEPSSKHQHKLVNRDLTKQSGSEMCKFLSTLTIDGVTRS 300

Db |||||

Qy 241 KLVLNCTARTELNVGIDFNWEPSSKHQHKLVNRDLTKQSGSEMCKFLSTLTIDGVTRS 300

Db |||||

Qy 301 DQGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGBRVRI PAKYLGYPVP 360

Db |||||

Qy 301 DQGLYTCAASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGBRVRI PAKYLGYPVP 360

Db |||||

Qy 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSDTGNVTYVILTNPISEKQSHVSVLVVYP 420

Db |||||

Qy 361 EIKWYKNGIPLESNHTIKAGHVLTIMEVSDTGNVTYVILTNPISEKQSHVSVLVVYP 420

Db |||||

Qy 421 PQIGKSLISPVDSYQYGTQTTLCTVYAI PPPHHIHWYQLEBECANESQAVSVTNPY 480

Db |||||

Qy 421 PQIGKSLISPVDSYQYGTQTTLCTVYAI PPPHHIHWYQLEBECANESQAVSVTNPY 480

Db |||||

Qy 481 PCEWRSVDFQGNKIEVNKNQFALIEGKNKTSTLVIQANVSALYKCEAVNKVGRGE 540

Db |||||

Qy 481 PCEWRSVDFQGNKIEVNKNQFALIEGKNKTSTLVIQANVSALYKCEAVNKVGRGE 540

Db |||||

Qy 541 RVISFHVTRGPEITLQDMQPTQESVSLMCTADRSFTENLTWYKLGQPLPIHVGLPT 600

Db |||||

Qy 541 RVISFHVTRGPEITLQDMQPTQESVSLMCTADRSFTENLTWYKLGQPLPIHVGLPT 600

Db |||||

Qy 601 PCKNLDLTKLNATMFSNSTNDILIMELKNASLDQGDYVCLADRKTKKRHCVRQLT 660

Db |||||

Qy 601 PCKNLDLTKLNATMFSNSTNDILIMELKNASLDQGDYVCLADRKTKKRHCVRQLT 660

Db |||||

Qy 661 VLERVAPITIGNLENQTTSGIESIEVSCASGNPPPOIMWFKDNETLVEDSGIVLKDGNR 720

Db |||||

Qy 661 VLERVAPITIGNLENQTTSGIESIEVSCASGNPPPOIMWFKDNETLVEDSGIVLKDGNR 720

Db |||||

Qy 721 NLTIIRVRKDEGLYTCQACSVLGCAKVEAFFIIEGAQKTNLEIILVGTAVIAMFFWL 780

Db |||||

Qy 721 NLTIIRVRKDEGLYTCQACSVLGCAKVEAFFIIEGAQKTNLEIILVGTAVIAMFFWL 780

Db |||||

Qy 781 LLVILITVRKANGGELTKGYLSIVMDPDELPIDEHCHERLPYDASKWEPFRDLKLGKPL 840

Db |||||

```
Db      781  LTVIILRTVCRANGGELKTCYLSIVMDPDELPLDEHCERLPYDASKWEFFDRDLKLGKPL 840
Qy      841  GRGAFQGVIBADAFIDKATCTRTAVAKMLKEGATHSEHRALMSELKILIHIGHILNVN 900
Db      841  GRGAFQGVIBADAFIDKATCTRTAVAKMLKEGATHSEHRALMSELKILIHIGHILNVN 900
Qy      901  LIGACTKPGGGLMVIIEFCKFNGLSTVLRKRNFEVYKTKGARFGQKDYVGAIPVDLK 960
Db      901  LIGACTKPGGGLMVIIEFCKFNGLSTVLRKRNFEVYKTKGARFGQKDYVGAIPVDLK 960
Qy      961  RLDSITSSQSSASSGFVEEKSLSVDEEBAPELDYKDFLTLEHLICYSQVAKGMEFLA 1020
Db      961  RLDSITSSQSSASSGFVEEKSLSVDEEBAPELDYKDFLTLEHLICYSQVAKGMEFLA 1020
Qy     1021  SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR 1080
Db     1021  SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR 1080
Qy     1081  VYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRLKEGTRMRAPDYTTPEMYQTML 1140
Db     1081  VYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRLKEGTRMRAPDYTTPEMYQTML 1140
Qy     1141  DCWHGEPSPRTPSELVEHLGNLLQANAQDGKDYIVLPISETLSMEEDSGLSPTSPTS 1200
Db     1141  DCWHGEPSPRTPSELVEHLGNLLQANAQDGKDYIVLPISETLSMEEDSGLSPTSPTS 1200
Qy     1201  CMEEVECDPKFHYDNTAGISQVLRNSKRSRPSVKTFFEDIPLPEPEVKVIPPDDNQDTS 1260
Db     1201  CMEEVECDPKFHYDNTAGISQVLRNSKRSRPSVKTFFEDIPLPEPEVKVIPPDDNQDTS 1260
Qy     1261  GMYLASELKTLEDRTKLSPFGMVPKSRVSESGNSQTSQYSGYHSDDTDTTVYS 1320
Db     1261  GMYLASELKTLEDRTKLSPFGMVPKSRVSESGNSQTSQYSGYHSDDTDTTVYS 1320
Qy     1321  SERAELLKLEIGVGTGTAQIQLQPDGTTLSPPV 1356
Db     1321  SERAELLKLEIGVGTGTAQIQLQPDGTTLSPPV 1356
```

## RESULT 3

```
US-10-995-561-905
; Sequence 905, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 905
; LENGTH: 1306
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-905
```

```
Query Match 95.9%; Score 6806; DB 6; Length 1306;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 1303; Conservative 2; Mismatches 1; Indels 50; Gaps 1;

Qy      1  MESKVLLAVALMVCVETRAASVGLPSVSLDPLRISIQKDLITIKANTTLQITCRGORDLD 60
Db      1  MQSKVLLAVALMVCVETRAASVGLPSVSLDPLRISIQKDLITIKANTTLQITCRGORDLD 60
Qy      61  WLMPNNGSGEQRVEVTECDGFCCKTLTIPIKIGNDGTAYKCFYRETDLASVIYVVD 120
Db      61  WLMPNNGSGEQRVEVTECDGFCCKTLTIPIKIGNDGTAYKCFYRETDLASVIYVVD 120
Qy     121  YRGPFIASVSDQHGVIYITENKNTVVIICLGSISNLNLSLCARYPEKRPVDPGNRISWD 180
```

```
Db     121  YRSPFIASVSDQHGVIYITENKNTVVIICLGSISNLNLSLCARYPEKRPVDPGNRISWD 180
Qy     181  SKGFTIPSYMTISYAGWVCEAKINDESYQSIYIYVVVGRIYDVVLSPSHGIELSVGE 240
Db     181  SKGFTIPSYMTISYAGWVCEAKINDESYQSIYIYVVVGRIYDVVLSPSHGIELSVGE 240
Qy     241  KLVNCTARTELNVGIDFNWEYPSKHQHKLVNRDLKTQSGSEMKKFLSTLITDGVTRS 300
Db     241  KLVNCTARTELNVGIDFNWEYPSKHQHKLVNRDLKTQSGSEMKKFLSTLITDGVTRS 300
Qy     301  DQGLYTCASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIPAKYLGYPPP 360
Db     301  DQGLYTCASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGERVRIPAKYLGYPPP 360
Qy     361  EIKWYKNGIPLSNHTIKAGHVLTIMEVSESDTGNVTIILTNPISEKQSHVSVLVVYP 420
Db     361  EIKWYKNGIPLSNHTIKAGHVLTIMEVSESDTGNVTIILTNPISEKQSHVSVLVVYP 420
Qy     421  PQIGEKSLISPVDSYQYGTQTLCTVYAIIPPHHIIHWYQLEECANEPQSVSTNYP 480
Db     421  PQIGEKSLISPVDSYQYGTQTLCTVYAIIPPHHIIHWYQLEECANEPQSVSTNYP 480
Qy     481  PCBEWSEVEDFOGKNIIEVKNQFALIEGKNKTSTLVIOAANVSALYKCEAVNKVGRGE 540
Db     481  PCBEWSEVEDFOGKNIIEVKNQFALIEGKNKTSTLVIOAANVSALYKCEAVNKVGRGE 540
Qy     541  RVISFHVTRGPEITLQPDQMPTEQSVSLWCTADRSTFENLTWYKLGQPLPIHVGELPT 600
Db     541  RVISFHVTRGPEITLQPDQMPTEQSVSLWCTADRSTFENLTWYKLGQPLPIHVGELPT 600
Qy     601  PUCKNLDTLWKLNATMFSNSTDILIMELKNASLQDQGYVCLAQDRKTKRRCVVRQLT 660
Db     601  PUCKNLDTLWKLNATMFSNSTDILIMELKNASLQDQGYVCLAQDRKTKRRCVVRQLT 660
Qy     661  VLERVAPTTIGNLENQTTIGESIEVSCASGNPPPIIMWFKNETLVEDSGIVLKQGNR 720
Db     661  VLERVAPTTIGNLENQTTIGESIEVSCASGNPPPIIMWFKNETLVEDSGIVLKQGNR 720
Qy     721  NLTIIRVRKEDGLYTCQACSVLGCACVKAFFIIEGAQEKTNLEIILVGTAVIAMPFWL 780
Db     721  NLTIIRVRKEDGLYTCQACSVLGCACVKAFFIIEGAQEKTNLEIILVGTAVIAMPFWL 780
Qy     781  LLVILIRTVKRANGELKTCYLSIVMDPDELPLDEHCERLPYDASKWEFFDRDLKLGKPL 840
Db     781  LLVILIRTVKRANGELKTCYLSIVMDPDELPLDEHCERLPYDASKWEFFDRDLKLGKPL 840
Qy     841  GRGAFQGVIEADAFIDKATCTRTAVAKMLKEGATHSEHRALMSELKILIHIGHILNVN 900
Db     841  GRGAFQGVIEADAFIDKATCTRTAVAKMLKEGATHSEHRALMSELKILIHIGHILNVN 900
Qy     901  LIGACTKPGGGLMVIIEFCKFNGLSTVLRKRNFEVYKTKGARFGQKDYVGAIPVDLK 960
Db     901  LIGACTKPGGGLMVIIEFCKFNGLSTVLRKRNFEVYKTKGARFGQKDYVGAIPVDLK 960
Qy     961  RLDSITSSQSSASSGFVEEKSLSVDEEBAPELDYKDFLTLEHLICYSQVAKGMEFLA 1020
Db     961  RLDSITSSQSSASSGFVEEKSLSVDEEBAPELDYKDFLTLEHLICYSQVAKGMEFLA 1020
Qy     1021  SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR 1080
Db     971  SRKCIHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGDARLPLKWMAPETIFDR 1030
Qy     1081  VYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRLKEGTRMRAPDYTTPEMYQTML 1140
Db     1031  VYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRLKEGTRMRAPDYTTPEMYQTML 1090
Qy     1141  DCWHGEPSPRTPSELVEHLGNLLQANAQDGKDYIVLPISETLSMEEDSGLSPTSPTS 1200
Db     1091  DCWHGEPSPRTPSELVEHLGNLLQANAQDGKDYIVLPISETLSMEEDSGLSPTSPTS 1150
Qy     1201  CMEEVECDPKFHYDNTAGISQVLRNSKRSRPSVKTFFEDIPLPEPEVKVIPPDDNQDTS 1260
```

Db 1151 CMEEVECDPKFHYDNTAGISQYLQNSKRSRPSVSKTFEDIDPLEEPEVKVIPPDDNQDS 1210  
Qy 1261 GVLASELTLERDTKLSFSFGMWPSKRESVASEGSGNOTSGYQSGYHSDDTDTTVYS 1320  
Db 1211 GVLASELTLERDTKLSFSFGMWPSKRESVASEGSGNOTSGYQSGYHSDDTDTTVYS 1270  
Qy 1321 SEAEALLKLEIGVGTGTAQIILQPDGTTLSPPV 1356  
Db 1271 SEAEALLKLEIGVGTGTAQIILQPDGTTLSPPV 1306

RESULT 4  
US-11-075-047A-4  
; Sequence 4, Application US/11075047A  
; Publication No. US20060030000A1  
; GENERAL INFORMATION:  
; APPLICANT: ALITALO, et al.  
; TITLE OF INVENTION: GROWTH FACTOR BINDING CONSTRUCTS MATERIALS AND METHODS  
; FILE REFERENCE: 28967/39700A  
; CURRENT APPLICATION NUMBER: US/11/075,047A  
; CURRENT FILING DATE: 2005-03-07  
; PRIOR APPLICATION NUMBER: US 60/550,907  
; PRIOR FILING DATE: 2004-03-07  
; NUMBER OF SEQ ID NOS: 128  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4  
; LENGTH: 764  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-075-047A-4

Query Match 56.7%; Score 4020; DB 7; Length 764;  
Best Local Similarity 100.0%; Pred. No. 9.9e-248; Indels 0; Gaps 0;  
Matches 764; Conservative 0; Mismatches 0;

Qy 1 MESKVLLAVALWLCVETRAASVGLPSVLDLPLSLQKDIILTIKANTTLQITCRGQDLD 60  
Db 1 MESKVLLAVALWLCVETRAASVGLPSVLDLPLSLQKDIILTIKANTTLQITCRGQDLD 60  
Qy 61 WLPNNQSGEQREVEVEGSDGLFCKTLTIPIKVGNDTGAYKCFYRETDLASVIVYVQD 120  
Db 61 WLPNNQSGEQREVEVEGSDGLFCKTLTIPIKVGNDTGAYKCFYRETDLASVIVYVQD 120  
Qy 121 YRSPFIASVSDQGVVYITENKNTVVIPLCGISNLNLSLCARYEKEKFPVDPGNRISWD 180  
Db 121 YRSPFIASVSDQGVVYITENKNTVVIPLCGISNLNLSLCARYEKEKFPVDPGNRISWD 180  
Qy 181 SKKGFTIPSYMISYAGMVFCEAKINDESYSQIMYIVVVGYRIYDVVLSPSHGIELSVGE 240  
Db 181 SKKGFTIPSYMISYAGMVFCEAKINDESYSQIMYIVVVGYRIYDVVLSPSHGIELSVGE 240  
Qy 241 KLVLNCTARTELNVGIDFNWEYSPSSKHQKLVNRDLKTQSGSEMCKFLSTLTIDGVTRS 300  
Db 241 KLVLNCTARTELNVGIDFNWEYSPSSKHQKLVNRDLKTQSGSEMCKFLSTLTIDGVTRS 300  
Qy 301 DQGLYTCASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGRVRIPAKILGYPPP 360  
Db 301 DQGLYTCASSGLMTKKNSTFVRVHEKPFVAFSGMESLVEATVGRVRIPAKILGYPPP 360  
Qy 361 EIKWYKNGIPLESNHTIKAGHVLTI MEVSERDGTGNVTILTNPISEKQSHVVSLVVYVP 420  
Db 361 EIKWYKNGIPLESNHTIKAGHVLTI MEVSERDGTGNVTILTNPISEKQSHVVSLVVYVP 420  
Qy 421 PQIGEKSLISPVDYSYQGTQTTLCTVYAI PPPHHIHWYQLEECANEPSSQAVSTNYP 480  
Db 421 PQIGEKSLISPVDYSYQGTQTTLCTVYAI PPPHHIHWYQLEECANEPSSQAVSTNYP 480  
Qy 481 PCEEWRSVDFQGNKIENKNOFALIEGKNTVSTLVIOANVSALYKCEAVNKVGRGE 540  
Db 481 PCEEWRSVDFQGNKIENKNOFALIEGKNTVSTLVIOANVSALYKCEAVNKVGRGE 540  
Qy 541 RVISFHVTRGPEITLQPDMPQTEQESVSLWCTADRTFFENLTWYKLGQPOLPIHVGSELPT 600

Db 541 RVISFHVTRGPEITLQPDMPQTEQESVSLWCTADRTFFENLTWYKLGQPOLPIHVGSELPT 600  
Qy 601 PVCKNLDTLWKLNATNFSNSTNDILIMELKNASLQDQGDYVCLADRKTKKRHCYVRQLT 660  
Db 601 PVCKNLDTLWKLNATNFSNSTNDILIMELKNASLQDQGDYVCLADRKTKKRHCYVRQLT 660  
Qy 661 VLERVAPTTIGNLENQTTISGESIEVSCASGNPPPOIMWFKDNETLVEDSGIVLKDGNR 720  
Db 661 VLERVAPTTIGNLENQTTISGESIEVSCASGNPPPOIMWFKDNETLVEDSGIVLKDGNR 720  
Qy 721 NLTIIRVRKEDGLYTCQACSVLGCACVKAFFIIEGAQEKTNLE 764  
Db 721 NLTIIRVRKEDGLYTCQACSVLGCACVKAFFIIEGAQEKTNLE 764

RESULT 5  
US-11-043-693-2  
; Sequence 2, Application US/11043693  
; Publication No. US20050281831A1  
; GENERAL INFORMATION:  
; APPLICANT: Davis-Smyth, Terri L.  
; APPLICANT: Chen, Helen H.  
; APPLICANT: Presta, Leonard  
; APPLICANT: Ferrara, Napoleone  
; TITLE OF INVENTION: NOVEL INHIBITORS OF VASCULAR ENDOTHELIAL  
; TITLE OF INVENTION: GROWTH FACTOR ACTIVITY, THEIR USES AND PROCESSES FOR THEIR  
; TITLE OF INVENTION: PRODUCTION  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dorsey & Whitney LLP  
; STREET: Four Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/11/043,693  
; FILING DATE: 26-Jan-2005  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/10/105,901  
; FILING DATE: 20-Mar-2002  
; APPLICATION NUMBER: 09/348,886  
; FILING DATE: 01-JUL-1999  
; APPLICATION NUMBER: US 08/643,839  
; FILING DATE: 07-MAY-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Richard F. Trecartin  
; REGISTRATION NUMBER: 31,801  
; REFERENCE/DOCKET NUMBER: A-63291-3/RFT/NBC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 398-3249  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 767 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-11-043-693-2

Query Match 56.6%; Score 4017; DB 7; Length 767;  
Best Local Similarity 99.9%; Pred. No. 1.5e-247;  
Matches 763; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESKVLLAVALWLCVETRAASVGLPSVLDLPLSLQKDIILTIKANTTLQITCRGQDLD 60

Db 1 MESKVLAVLWLCVETRAASVGLPSVSLDLPLRSIQKDLITIKANTTLQITCRQDLD 60  
Qy 61 WLWPNNGSGSRQVEVTECSGLFCKTLTIIPKVI GNDTGAYKCFYRETDLASVIYVVD 120  
Db 61 WLWPNNGSGSRQVEVTECSGLFCKTLTIIPKVI GNDTGAYKCFYRETDLASVIYVVD 120  
Qy 121 YRSPFIASVSDQHGVIYITENKNTVITPCIGSISLNLSVCARYPEKRFVPGDNRISWD 180  
Db 121 YRSPFIASVSDQHGVIYITENKNTVITPCIGSISLNLSVCARYPEKRFVPGDNRISWD 180  
Qy 181 SKGFTTIPSYMISYAGWVFCFAKINDESYQSIMIYVVGRIYDVVLSHSHGIELSVGE 240  
Db 181 SKGFTTIPSYMISYAGWVFCFAKINDESYQSIMIYVVGRIYDVVLSHSHGIELSVGE 240  
Qy 241 KLVLNCTARTELNAGIDFNWEPSSKHQKHLVNRDLKTQSGSEMCKFLSTLTIDGVTRS 300  
Db 241 KLVLNCTARTELNAGIDFNWEPSSKHQKHLVNRDLKTQSGSEMCKFLSTLTIDGVTRS 300  
Qy 301 DOGLYTCAASGLMTKKNSTFVRVHEKPFVAFGSGMESLVEATVGERVRIIPAKYLGYP 360  
Db 301 DOGLYTCAASGLMTKKNSTFVRVHEKPFVAFGSGMESLVEATVGERVRIIPAKYLGYP 360  
Qy 361 EIKWYKNGIPLSNHTIKAGHVLTIMEVSRDTCNTVILTNPISEKQSHVSLVYVP 420  
Db 361 EIKWYKNGIPLSNHTIKAGHVLTIMEVSRDTCNTVILTNPISEKQSHVSLVYVP 420  
Qy 421 POIGKSLISPVDSYQYGTQTTLCTVVAIPPPHHIHWYQLEBECANEPQAVSVTPNY 480  
Db 421 POIGKSLISPVDSYQYGTQTTLCTVVAIPPPHHIHWYQLEBECANEPQAVSVTPNY 480  
Qy 481 PCBEWRSVEDFQGNKEVKNQFALLEGKNTVSTLVIQANVSALYKCAVNVKVRGE 540  
Db 481 PCBEWRSVEDFQGNKEVKNQFALLEGKNTVSTLVIQANVSALYKCAVNVKVRGE 540  
Qy 541 RVISFHVTRGPEITLQDPMQTEQSVSLWCTADRSTFENLTWYKLGQPLPIHVGELPT 600  
Db 541 RVISFHVTRGPEITLQDPMQTEQSVSLWCTADRSTFENLTWYKLGQPLPIHVGELPT 600  
Qy 601 PVCKNLDLTKLNATMFSNSTNDILIMELKNASLDQDGYVCLAQDRKTKRHCVRQLT 660  
Db 601 PVCKNLDLTKLNATMFSNSTNDILIMELKNASLDQDGYVCLAQDRKTKRHCVRQLT 660  
Qy 661 VLERVAPITGNLENQTTSGESIEVSTAGNPPQIMPFKONETLVEDSGIVLKDGNR 720  
Db 661 VLERVAPITGNLENQTTSGESIEVSTAGNPPQIMPFKONETLVEDSGIVLKDGNR 720  
Qy 721 NLTRRRVKEDEGLYTQACSVLGCACVKAFFIIEGAQKTNLE 764  
Db 721 NLTRRRVKEDEGLYTQACSVLGCACVKAFFIIEGAQKTNLD 764

## RESULT 6

US-11-104-110-8  
; Sequence 8, Application US/11104110  
; Publication No. US2006002916A1  
; GENERAL INFORMATION:  
; APPLICANT: Nguyen, Jack  
; TITLE OF INVENTION: CLEAVAGE OF VEGF AND VEGF RECEPTOR BY WILDTYPE AND MUTANT MT-SPI  
; FILE REFERENCE: 25840-502  
; CURRENT FILING DATE: 2005-04-12  
; PRIOR FILING DATE: 2004-04-12  
; PRIOR FILING DATE: 2003-10-02  
; PRIOR FILING DATE: 2002-10-02  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 8  
; LENGTH: 764

; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-104-110-8

Query Match : 56.5%; Score 4012; DB 7; Length 764;  
Best Local Similarity 99.9%; Pred. No. 3.2e-247;  
Matches 762; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ESKVLLAVLWLCVETRAASVGLPSVSLDLPLRSIQKDLITIKANTTLQITCRQDLD 61  
Db 2 QSKVLLAVLWLCVETRAASVGLPSVSLDLPLRSIQKDLITIKANTTLQITCRQDLD 61  
Qy 62 LWPNNQSGSRQVEVTECSGLFCKTLTIIPKVI GNDTGAYKCFYRETDLASVIYVVD 121  
Db 62 LWPNNQSGSRQVEVTECSGLFCKTLTIIPKVI GNDTGAYKCFYRETDLASVIYVVD 121  
Qy 122 RSPFIASVSDQHGVIYITENKNTVITPCIGSISLNLSVCARYPEKRFVPGDNRISWD 181  
Db 122 RSPFIASVSDQHGVIYITENKNTVITPCIGSISLNLSVCARYPEKRFVPGDNRISWD 181  
Qy 182 KGFTTIPSYMISYAGWVFCFAKINDESYQSIMIYVVGRIYDVVLSHSHGIELSVGE 241  
Db 182 KGFTTIPSYMISYAGWVFCFAKINDESYQSIMIYVVGRIYDVVLSHSHGIELSVGE 241  
Qy 242 LVLNCTARTELNAGIDFNWEPSSKHQKHLVNRDLKTQSGSEMCKFLSTLTIDGVTRS 301  
Db 242 LVLNCTARTELNAGIDFNWEPSSKHQKHLVNRDLKTQSGSEMCKFLSTLTIDGVTRS 301  
Qy 302 QGLYTCAASGLMTKKNSTFVRVHEKPFVAFGSGMESLVEATVGERVRIIPAKYLGYP 361  
Db 302 QGLYTCAASGLMTKKNSTFVRVHEKPFVAFGSGMESLVEATVGERVRIIPAKYLGYP 361  
Qy 362 IKWYKNGIPLSNHTIKAGHVLTIMEVSRDTCNTVILTNPISEKQSHVSLVYVP 421  
Db 362 IKWYKNGIPLSNHTIKAGHVLTIMEVSRDTCNTVILTNPISEKQSHVSLVYVP 421  
Qy 422 QIGKSLISPVDSYQYGTQTTLCTVVAIPPPHHIHWYQLEBECANEPQAVSVTPNY 481  
Db 422 QIGKSLISPVDSYQYGTQTTLCTVVAIPPPHHIHWYQLEBECANEPQAVSVTPNY 481  
Qy 482 CBEWRSVEDFQGNKEVKNQFALLEGKNTVSTLVIQANVSALYKCAVNVKVRGE 541  
Db 482 CBEWRSVEDFQGNKEVKNQFALLEGKNTVSTLVIQANVSALYKCAVNVKVRGE 541  
Qy 542 VISFHVTRGPEITLQDPMQTEQSVSLWCTADRSTFENLTWYKLGQPLPIHVGELPT 601  
Db 542 VISFHVTRGPEITLQDPMQTEQSVSLWCTADRSTFENLTWYKLGQPLPIHVGELPT 601  
Qy 602 VCKNLDLTKLNATMFSNSTNDILIMELKNASLDQDGYVCLAQDRKTKRHCVRQLT 661  
Db 602 VCKNLDLTKLNATMFSNSTNDILIMELKNASLDQDGYVCLAQDRKTKRHCVRQLT 661  
Qy 662 LERVAPITGNLENQTTSGESIEVSTAGNPPQIMPFKONETLVEDSGIVLKDGNR 721  
Db 662 LERVAPITGNLENQTTSGESIEVSTAGNPPQIMPFKONETLVEDSGIVLKDGNR 721  
Qy 722 LTIRRVKEDGLYTQACSVLGCACVKAFFIIEGAQKTNLE 764  
Db 722 LTIRRVKEDGLYTQACSVLGCACVKAFFIIEGAQKTNLE 764

## RESULT 7

US-11-104-111-28  
; Sequence 28, Application US/11104111  
; Publication No. US20060024289A1  
; GENERAL INFORMATION:  
; APPLICANT: Waugh Ruggles, Sandra  
; TITLE OF INVENTION: CLEAVAGE OF VEGF AND VEGF RECEPTOR BY WILD-TYPE AND MUTANT  
; FILE REFERENCE: 25840-503  
; CURRENT FILING DATE: 2005-04-12

; PRIOR APPLICATION NUMBER: 60/561,671  
; PRIOR FILING DATE: 2004-04-12  
; PRIOR APPLICATION NUMBER: 10/677,977  
; PRIOR FILING DATE: 2003-10-02  
; PRIOR APPLICATION NUMBER: 60/415,388  
; PRIOR FILING DATE: 2002-10-02  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 28  
; LENGTH: 764  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-104-111-28

Query Match 56.5%; Score 4012; DB 7; Length 764;  
Best Local Similarity 99.9%; Pred. No. 3.2e-247;  
Matches 762; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 ESKVLLAVALWLCVETRAASVGLPSVSLDPLRLSIQKDILTIKANTTLQITCRGQRLDW 61  
Db :  
Qy 62 LWPNNQSGSQREVEVTECSGDLFKTLTIPKVIQNDTGAYKCFYRETDLASVIYVYQDY 121  
Db 62 LWPNNQSGSQREVEVTECSGDLFKTLTIPKVIQNDTGAYKCFYRETDLASVIYVYQDY 121  
Qy 122 RSPFIASVSDQGVVYITENKNTVPCIGSISNLNVSICARYPEKRFVDPDGNRISWDS 181  
Db 122 RSPFIASVSDQGVVYITENKNTVPCIGSISNLNVSICARYPEKRFVDPDGNRISWDS 181  
Qy 182 KGFTIPSYMISYAGMVFCEAKINDESYQSIMYIVVVGYYRIYDVVLSPSHGHISLVSGEK 241  
Db 182 KGFTIPSYMISYAGMVFCEAKINDESYQSIMYIVVVGYYRIYDVVLSPSHGHISLVSGEK 241  
Qy 242 LVNLCTARTLNIGIDFNWYPSKQHKLVNRDLKTQSGSEMKFLSTLTIDGVTRSD 301  
Db 242 LVNLCTARTLNIGIDFNWYPSKQHKLVNRDLKTQSGSEMKFLSTLTIDGVTRSD 301  
Qy 302 QGLYTCAASGLMTKKNSTFVRVHEKPFVAFGSMESLVEATVGERVRIIPAKYLGVPPE 361  
Db 302 QGLYTCAASGLMTKKNSTFVRVHEKPFVAFGSMESLVEATVGERVRIIPAKYLGVPPE 361  
Qy 362 IKWYKNGIPILESNHTIKAGHVLTIMEVSRDGTNYTILTNPISEKQSHVSVLSVYVVP 421  
Db 362 IKWYKNGIPILESNHTIKAGHVLTIMEVSRDGTNYTILTNPISEKQSHVSVLSVYVVP 421  
Qy 422 QIGKSLISPVDSYQYGTOTLTCTVVAIPPHHHIHWYOLEBECANEPQAVSVTNYP 481  
Db 422 QIGKSLISPVDSYQYGTOTLTCTVVAIPPHHHIHWYOLEBECANEPQAVSVTNYP 481  
Qy 482 CEEWRSVEDFOGKNIEVKNQFALIEGKNKTVSTLVIQANYSALYKCEAVNKVGRGER 541  
Db 482 CEEWRSVEDFOGKNIEVKNQFALIEGKNKTVSTLVIQANYSALYKCEAVNKVGRGER 541  
Qy 542 VISFHVTRGPEITLQPDQMTEQESVSLWCTADRSFENLTWYKLGPOPLPIHVGBLPTP 601  
Db 542 VISFHVTRGPEITLQPDQMTEQESVSLWCTADRSFENLTWYKLGPOPLPIHVGBLPTP 601  
Qy 602 VCKNLDLTLWKNATMFSNSTNDILIMELKNASLQDQGYVCLAQDRKTKRCHCVRLTV 661  
Db 602 VCKNLDLTLWKNATMFSNSTNDILIMELKNASLQDQGYVCLAQDRKTKRCHCVRLTV 661  
Qy 662 LERVAPITITLANOTTSIGESIEVSTASGNPPQIMWFKNETLVEDSGIIVKGNRN 721  
Db 662 LERVAPITITLANOTTSIGESIEVSTASGNPPQIMWFKNETLVEDSGIIVKGNRN 721  
Qy 722 LTIIRRVKDEGLYTCOACSVLGCCKVAFFIIEGAQKTNLE 764  
Db 722 LTIIRRVKDEGLYTCOACSVLGCCKVAFFIIEGAQKTNLE 764

RESULT 8  
US-11-104-110-9

; Sequence 9, Application US/11104110  
; Publication No. US20060002916A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruggles, Sandra  
; APPLICANT: Nguyen, Jack  
; TITLE OF INVENTION: CLEAVAGE OF VEGF AND VEGF RECEPTOR BY WILDTYPE AND MUTANT MT-SP1  
; FILE REFERENCE: 25840-502  
; CURRENT APPLICATION NUMBER: US/11/104,110  
; CURRENT FILING DATE: 2005-04-12  
; PRIOR APPLICATION NUMBER: 60/561,720  
; PRIOR FILING DATE: 2004-04-12  
; PRIOR APPLICATION NUMBER: 10/677,977  
; PRIOR FILING DATE: 2003-10-02  
; PRIOR APPLICATION NUMBER: 60/415,388  
; PRIOR FILING DATE: 2002-10-02  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 9  
; LENGTH: 592  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-104-110-9

Query Match 43.3%; Score 3075; DB 7; Length 592;  
Best Local Similarity 100.0%; Pred. No. 7.9e-188;  
Matches 592; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 765 IILVGTAVIAMPFWLLVIIILRTVVRANGGELKTGYLSIVMDPDELPLDHCERLPYDA 824  
Db 1 IILVGTAVIAMPFWLLVIIILRTVVRANGGELKTGYLSIVMDPDELPLDHCERLPYDA 60  
Qy 825 SKWEFPRDLKLGKPLGRGAFQVIEADAFIDKATCTRTAVAKMLKEGATSEHRAALMS 884  
Db 61 SKWEFPRDLKLGKPLGRGAFQVIEADAFIDKATCTRTAVAKMLKEGATSEHRAALMS 120  
Qy 885 ELKILIHGHLLNVNLLGACTKPGGFLMVIVVEFCFKNLSTYLRSKRNEFVYKTKGAR 944  
Db 121 ELKILIHGHLLNVNLLGACTKPGGFLMVIVVEFCFKNLSTYLRSKRNEFVYKTKGAR 180  
Qy 945 FROGKDYVGAIPVDLKRRLDSITSSQSSASSGVEEKSISLSDVEEAPEDLYKDFLTLEH 1004  
Db 181 FROGKDYVGAIPVDLKRRLDSITSSQSSASSGVEEKSISLSDVEEAPEDLYKDFLTLEH 240  
Qy 1005 LICYSFOVAKGMEFLASRCKIHRDLAARNLILLSEKNVVKICDFGLARDIYKDPDYVRKGD 1064  
Db 241 LICYSFOVAKGMEFLASRCKIHRDLAARNLILLSEKNVVKICDFGLARDIYKDPDYVRKGD 300  
Qy 1065 ARPLKXMAPETIFDRVYTIQSDVWSFGVLLWEIFSLGASPYPGVKIDEEFCRRLKEGTR 1124  
Db 301 ARPLKXMAPETIFDRVYTIQSDVWSFGVLLWEIFSLGASPYPGVKIDEEFCRRLKEGTR 360  
Qy 1125 MRAPDYTTTPEMTQMLDCHWGEPSPQRTFSELVEHGNLLQANAQDQKDIYVLPISL 1184  
Db 361 MRAPDYTTTPEMTQMLDCHWGEPSPQRTFSELVEHGNLLQANAQDQKDIYVLPISL 420  
Qy 1185 SMEEDSGLSIPSPVSCMEEEVECDPKFHYDNTAGISQYLSNKRKRSPVSVKTFEDIP 1244  
Db 421 SMEEDSGLSIPSPVSCMEEEVECDPKFHYDNTAGISQYLSNKRKRSPVSVKTFEDIP 480  
Qy 1245 ESEPEVKVIFDDNQTDGSMVLASEELKTLEDRTKLSFSGMVPKSGRESVASGSGNQTSG 1304  
Db 481 ESEPEVKVIFDDNQTDGSMVLASEELKTLEDRTKLSFSGMVPKSGRESVASGSGNQTSG 540  
Qy 1305 YQSGYHSDDDTTTVYSSEAEALLKLIIEIGVQTGSTAQIILQPDGTTLSPPV 1356  
Db 541 YQSGYHSDDDTTTVYSSEAEALLKLIIEIGVQTGSTAQIILQPDGTTLSPPV 592

RESULT 9  
US-11-104-111-29  
; Sequence 29, Application US/11104111  
; Publication No. US20060024289A1  
; GENERAL INFORMATION:

; APPLICANT: Waugh Ruggles, Sandra  
; APPLICANT: Nguyen, Jack  
; TITLE OF INVENTION: CLEAVAGE OF VEGF AND VEGF RECEPTOR BY WILD-TYPE AND MUTANT  
; FILE OF INVENTION: PROTEASES  
; FILE REFERENCE: 25840-503  
; CURRENT FILING DATE: 2005-04-12  
; PRIOR APPLICATION NUMBER: 60/561,671  
; PRIOR FILING DATE: 2004-04-12  
; PRIOR APPLICATION NUMBER: 10/677,977  
; PRIOR FILING DATE: 2003-10-02  
; PRIOR APPLICATION NUMBER: 60/415,388  
; PRIOR FILING DATE: 2002-10-02  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 29  
; LENGTH: 592  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-104-111-29

Query Match 43.3%; Score 3075; DB 7; Length 592;  
Best Local Similarity 100.0%; Pred. No. 7.9e-188;  
Matches 592; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 765 IILVGTAVTAMFPLMLLILRTVKANGELKTGYLSIVMDPDELPLDHCERLPYDA 824  
DB 1 IILVGTAVTAMFPLMLLILRTVKANGELKTGYLSIVMDPDELPLDHCERLPYDA 60  
QY 825 SKWEFPRDLKGLKPLGRGAGQVIEADAFIDKATCTRTVAVKMLKEGATHSEHRALMS 884  
DB 61 SKWEFPRDLKGLKPLGRGAGQVIEADAFIDKATCTRTVAVKMLKEGATHSEHRALMS 120  
QY 885 ELKILIHGHHLNVNLLGACTKPGPLMVI VEFCKFGNLSYLSKRNEFPVYKTKGAR 944  
DB 121 ELKILIHGHHLNVNLLGACTKPGPLMVI VEFCKFGNLSYLSKRNEFPVYKTKGAR 180  
QY 945 FROGKDYVGAIPVDLKRRLDSITSSQSSASSGFFVEEKSLSDVVEEAPEDLYKDFLTLEH 1004  
DB 181 FROGKDYVGAIPVDLKRRLDSITSSQSSASSGFFVEEKSLSDVVEEAPEDLYKDFLTLEH 240  
QY 1005 LICYSFOVAGMEFLASKKICHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGD 1064  
DB 241 LICYSFOVAGMEFLASKKICHRDLAARNILLSEKNVVKICDFGLARDIYKDPDYVRKGD 300  
QY 1065 ARPLKMWAPETIFDRVYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRLKEGTR 1124  
DB 301 ARPLKMWAPETIFDRVYTIQSDVMSFGVLLWEIFSLGASPYGVKIDEEFCRLKEGTR 360  
QY 1125 MRAPDYTPPEMYQTMDCWHGEPSSQRTFSELVHGLNLLQANAQQDGKDYIVLPISETL 1184  
DB 361 MRAPDYTPPEMYQTMDCWHGEPSSQRTFSELVHGLNLLQANAQQDGKDYIVLPISETL 420  
QY 1185 SMEEDSGLSLPTSPVSCWEEVEECVDPKPHYDNTAGISQYLNKRSRPSVKTFFEDIPL 1244  
DB 421 SMEEDSGLSLPTSPVSCWEEVEECVDPKPHYDNTAGISQYLNKRSRPSVKTFFEDIPL 480  
QY 1245 EEPEVKVIPDNDQSDGMVLASBELKTLDRTKLSPSFGGMPKSGRESVASGSGNQTSG 1304  
DB 481 EEPEVKVIPDNDQSDGMVLASBELKTLDRTKLSPSFGGMPKSGRESVASGSGNQTSG 540  
QY 1305 YQSGYHSDDDTTTYSSEAEELKLIEIGVQTGSTAQIQLPDSGTTLSSPPV 1356  
DB 541 YQSGYHSDDDTTTYSSEAEELKLIEIGVQTGSTAQIQLPDSGTTLSSPPV 592

## RESULT 10

US-11-043-693-34  
; Sequence 34, Application US/11043693  
; Publication No. US20050281831A1  
; GENERAL INFORMATION:  
; APPLICANT: Davis-Smyth, Terri L.  
; APPLICANT: Chen, Helen H.

; APPLICANT: Presta, Leonard  
; APPLICANT: Ferrara, Napoleone  
; TITLE OF INVENTION: NOVEL INHIBITORS OF VASCULAR ENDOTHELIAL  
; TITLE OF INVENTION: GROWTH FACTOR ACTIVITY, THEIR USES AND PROCESSES FOR THEIR  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dorsey & Whitney LLP  
; STREET: Four Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/11/043,693  
; FILING DATE: 26-Jan-2005  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/10/105,901  
; FILING DATE: 20-Mar-2002  
; APPLICATION NUMBER: 09/348,886  
; FILING DATE: 01-JUL-1999  
; APPLICATION NUMBER: US 08/643,839  
; FILING DATE: 07-MAY-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Richard F. Trecartin  
; REGISTRATION NUMBER: 31,801  
; REFERENCE/DOCKET NUMBER: A-63291-3/RFT/NBC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 398-3249  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1368 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-11-043-693-34

Query Match 39.4%; Score 2977.5; DB 7; Length 1368;  
Best Local Similarity 44.6%; Pred. No. 1e-169;  
Matches 615; Conservative 204; Mismatches 472; Indels 89; Gaps 29;  
QY 1 MESKVLVALVMLCVETRAASVGLPSVSLDPLRLSI-QKDILTIKANTTLQITCRGQDL 59  
DB 1 MORGAAALCLRLWLCGLLDGLVS--GYSMTPTLSITTESHVIDTGDLSLSICRGQHL 58  
QY 60 DWLWPNQSG-----SEORVEVTECSDEL-----PCKLTIPKVIQNDTGAKCFYR-- 106  
DB 59 EAWPQAQAPATGDKDSEDGTGVVRC-EGTDARPYCKVLLLEHVANDTGSVCYKYI 117  
QY 107 ----ETDLASVIVVQDYRSPFIASVDQHGVIYITENKNTKWTWIPCLGSISNLNLSLC 162  
DB 118 KARIETGTAASIVIFISDTGRPFVEMYRIPEIHHTEGRE--LVIPCRVTSNIITVL- 174  
QY 163 ARYPEKRFVDPGNRISWDSKKGFTIPSYMISVAGMVFCRAKINDESYQSIYIVVVGYR 222  
DB 175 KKEPDLTLPDGKRIIWDNRKGFIIISNATYKEIGLUTCEATVNGHLYKT-NYLTHRTGNE 233  
QY 223 IYDVVLSPSHGTELSVGEKLVNLTARTLNVCIDFNWEYPPSSKHQHKLVNRLDKTQSG 282  
DB 234 LYDIQLLPKRSLELLVGEKLVNLTCTVWAEFNSGVTDFWDYPGQAERGKWPERRSQOTH 293  
QY 283 SEMKKFLSTLTIDGVTRSDOGLYTCASSGLMTKKNSTFVRVHEKPPVAFGSGMESLVEA 342  
DB 294 TELS-----SILTIHNSQHDLSYVCKANNNGIQRFRESTEVRVHENPFISVLEWLGPILEA 350











Qy	928	LRSKRNEFVYKTKG-----ARPRQGKDYGAI PVDLKR--RLDSITSSQSASSSGFVEEK	981
Dd	939	LRAKRDAFSPCAEKSPEQRGRFRA---MVELARLDRRPGSSDRVLFAFPFKTEGGARRA	995
Qy	982	SLSDVEEBEAPEDLYKDFLTLEHLICYSFOVAKGMELASRKCIHRDLAARNILLSEKNV	1041
Dd	996	S-----PQDEA-EDLWLSPLTWEDLVCSFQVARGMEFLASRKCIHRDLAARNILLSDDV	1050
Qy	1042	VKICDGLARDIYKDPDYVRKGDKARLPKWMAPEYIFDRVVTIQSDVMSFGVLLWEIPSL	1101
Dd	1051	VKICDGLARDIYKDPDYVRKGSARLPKWMAPEYIFDVVVTTQSDVMSFGVLLWEIPSL	1110
Qy	1102	GASPYGVKIDBEFCRRLKGTMRADPTTYPMYQTMLDCWHCEGPQRPFTFSELVEHIG	1161
Dd	1111	GASPYGVGVINEBFCQLRDTGTRMAPELATPAIRRMILNLCWSGDPKARPAPFSELVEILG	1170
Qy	1162	NLLQANAQDGKDIVLPISETLSMBEDSGLSLTPSPVSCMBEEVCDP-----KPH	1211
Dd	1171	DLLOCRGLQEIEEVCMAPRS-SQSSSEGSPSQVSTMALHTAQADAESPPSLQRHSLAAR	1222
Qy	1214	YDNTAGISQYLONSKRKRPSVKTFEDIPLBEPVKKVIPDNDQTDGSMVLASELKTLTLE	1277
Dd	1230	YNNWVPFGCLARGAETRGSRMKTPFEFFM-TPTTYKGSVDNQTDGSMVLASEFEQIE	1288
Qy	1274	DRTKLSPSF 1282	
Dd	1289	SRHQESGF 1297	
 RESULT 15			
US-11-109-156-23			
; Sequence 23, Application US/11109156			
; Publication No. US20050250144A1			
; GENERAL INFORMATION:			
; APPLICANT: Toshio Ota			
; APPLICANT: Takao Isogai			
; APPLICANT: Tetsuo Nishikawa			
; APPLICANT: Koji Hayashi			
; APPLICANT: Kaoru Otsuka			
; APPLICANT: Jun-Ichi Yamamoto			
; APPLICANT: Shizuko Ishii			
; APPLICANT: Tomoyasu Sugiyama			
; APPLICANT: Ai Wakamatsu			
; APPLICANT: Keiichi Nagai			
; APPLICANT: Tetsuji Otsuki			
; APPLICANT: Shin-Ichi Funahashi			
; APPLICANT: Chiaki Senoo			
; APPLICANT: Jun-Ichi Nezu			
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN			
; TITLE OF INVENTION: PHOSPHATASE			
; FILE REFERENCE: 06501-099002			
; CURRENT APPLICATION NUMBER: US/11/109,156			
; CURRENT FILING DATE: 2005-04-19			
; PRIOR APPLICATION NUMBER: US/10/060,065			
; PRIOR FILING DATE: 2002-01-29			
; PRIOR APPLICATION NUMBER: PCT/JPO0/05061			
; PRIOR FILING DATE: 2000-07-28			
; PRIOR APPLICATION NUMBER: US 60/159,590			
; PRIOR FILING DATE: 1999-10-18			
; PRIOR APPLICATION NUMBER: US 60/183,322			
; PRIOR FILING DATE: 2000-02-17			
; PRIOR APPLICATION NUMBER: JP 11-248036			
; PRIOR FILING DATE: 1999-07-29			
; PRIOR APPLICATION NUMBER: JP 2000-118776			
; PRIOR FILING DATE: 2000-01-11			
; PRIOR APPLICATION NUMBER: JP 2000-183767			
; PRIOR FILING DATE: 2000-05-02			
; PRIOR APPLICATION NUMBER: JP 2000-241899			
; PRIOR FILING DATE: 2000-06-09			
; NUMBER OF SEQ ID NOS: 43			
; SOFTWARE: Patentin Ver. 2.0			
; SEQ ID NO 23			
; LENGTH: 1338			

Query Match 38.8%; Score 2751.5; DB 7; Length 1298;  
Best Local Similarity 44.9%; Pred. No. 8.1e-167;  
Matches 597; Conservative 195; Mismatches 458; Indels 79; Gaps 24;

US-11-075-047A-6

1 MESKVLAVALLVCVETRAASVGLSPSVSLDLRLSLQIDILTIKANTTQITCRQDRLD 60  
1 MQRGAALCLRLMLCGLLDGLVS--GYSMTPTLNTIETHSVIDTGDLSLSICRGQHPLE 58  
61 WLPNNQSG-----SEQRVEVTECSDEL-----FKTLTIPKIVGNTDGTAYKCFYR--- 106  
59 WAMPGAQEA PATGDKDSEDTGVVRDC-EGTDARPYCKVLLHEVHANDTGSVVCYKYIK 117  
107 ---ETDLASVIVYVVDYSPPTASVDQHGYYITENKNTVPIPCLGISINLSICA 163  
118 ARTEGTTAASSVYVVRDFPQPFINKPDT-----LLVNRKAMWVPCLVSLPGLNVL-- 169  
164 RYPERKFPDGNRISWDSKKGITIPSYMISYAGWVFCFAKINDESYQSIIMTVVVVGYRI 223  
170 RSQSSVLMFDQGEVWDDRRGMLVSTPLLDALYLQCTTGWCDQDFLSNPLVHITGNEL 229  
224 YDVVLSPSHGIELSGEKLVLNCTARTELNVGIDFNWBPYSPKHOHKLNVRLDKLTQSGS 283  
230 YDIQLPRKSLLELVGKLVNCTVMAEFNSGVTFDWDYPGKQABRGKWPERRSQOHT 289  
284 EMKFLSTLTIDGVTSDOGLYTCAASSGLMTKNSSTFVRVHEKPPVAFGSGMESLEAT 343  
290 ELS---SILTIHNSQHDIGSVCKANNGIORFRESTEVI VHENPFI SVEWLKGPILEAT 346  
344 VG-ERVRIIPAKLYGPPPEIKVYKXGIPLESNHTIKAGHLVTIMEVSRDGTGNTVILT 402  
347 AGDELVKLPVLAAYPPPEFQYKQKALSGRH---SPHALVKKEVTEASTGTTTLALWN 403  
403 PISKEQSHVSLVWVVPVPPGIEKSLISVSDSYQYGTQTTLCTVYAIAPPHHIIHWYQL 462  
404 SAAGLRNRISLELVNVPPIHEKEASSP-SIYSRHSRQALTCTAYGVPLPLSIQHWHP 462  
463 BEEC---ANEPSQNVSTWNPY-CEWRSVEDFQGNKIEVKNQFALIEGKNTVSTLV 518  
463 WTPCKMFAQRSLRRRQQQDLMPQCRDRAVTTQDAVNPIESLDTWTEFVEGKNTVSKLV 522  
519 IQANVSALYKCEAVNKVGRGSRVLSFHYTRGPE---ITLQDPMQPTQESVSLWCTADR 575  
523 IQANVSAMTKCVSNKVGQDERLYFYVYTTIPDGFTIESPSEBELGGQPVLLSCQADS 582  
576 STPENITWYKLGPPQLPIHVGELTPVCKNLDTLWKLNATMFSNSTNDI-----LIM 627  
583 KYEHLRWYRLNLSLTHDAHGNPLLLDCKQV---HLFATPLAASLEEVAPGARHATLSL 638  
628 ELKNASLDQDGYVCLAQDRKTKRHCVVRQVTLVLERVAPITIGNLEQNTTSIGESIYS 687  
639 SIPRVAPEHEGYVCEVQDRRSHDKHCKKYLVSQALEAPRLTQLTDLNVNDSLEMQ 698  
688 CTASGNPPQIMWFKDNETLVEDSGVLVDKGNRNTIRVRKEDGLYTCOACSVLGCAK 747  
699 CLVAGAHAPSIWYKDERLEEKSGVDLADSNOKLSIORVEEDAGRYLCSVCAKGCVN 758  
748 VEAFFITGAQKTNLEIILVGTAVIAMFFWLLLVILRTVKRANGGELKTGYLSIVMD 807  
759 SSASVAVESGEDKGSNEIVLGTGVIAVFFWVLLLLIFCNWRRAPHADIKTYLSIIMD 818  
808 PDELPLDECHERLPYDASKWEPFRDLKLGKPLGRGAFQGVIEADAFGDKTATCTRTAV 867  
819 PGEVPLEEQCEYLSYDASQWEPFRERHLGRVLGYGAFKGVVEASAFGIHKSSCDTAV 878  
868 KMLKEGATHSEHALMSELKILIHGHLLNVNVLGACTKPGCPPLMWIVECKGNLSTY 927  
879 KMLKEGATASEHALMSELKILIHGHLLNVNVLGACTKPGCPPLMWIVECKGNLSTY 938

Query Match 38.8%; Score 2751.5; DB 7; Length 1298;  
Best Local Similarity 44.9%; Pred. No. 8.1e-167;  
Matches 597; Conservative 195; Mismatches 458; Indels 79; Gaps 24;

US-11-075-047A-6

1 MESKVLAVALLVCVETRAASVGLSPSVSLDLRLSLQIDILTIKANTTQITCRQDRLD 60  
1 MQRGAALCLRLMLCGLLDGLVS--GYSMTPTLNTIETHSVIDTGDLSLSICRGQHPLE 58  
61 WLPNNQSG-----SEQRVEVTECSDEL-----FKTLTIPKIVGNTDGTAYKCFYR--- 106  
59 WAMPGAQEA PATGDKDSEDTGVVRDC-EGTDARPYCKVLLHEVHANDTGSVVCYKYIK 117  
107 ---ETDLASVIVYVVDYSPPTASVDQHGYYITENKNTVPIPCLGISINLSICA 163  
118 ARTEGTTAASSVYVVRDFPQPFINKPDT-----LLVNRKAMWVPCLVSLPGLNVL-- 169  
164 RYPERKFPDGNRISWDSKKGITIPSYMISYAGWVFCFAKINDESYQSIIMTVVVVGYRI 223  
170 RSQSSVLMFDQGEVWDDRRGMLVSTPLLDALYLQCTTGWCDQDFLSNPLVHITGNEL 229  
224 YDVVLSPSHGIELSGEKLVLNCTARTELNVGIDFNWBPYSPKHOHKLNVRLDKLTQSGS 283  
230 YDIQLPRKSLLELVGKLVNCTVMAEFNSGVTFDWDYPGKQABRGKWPERRSQOHT 289  
284 EMKFLSTLTIDGVTSDOGLYTCAASSGLMTKNSSTFVRVHEKPPVAFGSGMESLEAT 343  
290 ELS---SILTIHNSQHDIGSVCKANNGIORFRESTEVI VHENPFI SVEWLKGPILEAT 346  
344 VG-ERVRIIPAKLYGPPPEIKVYKXGIPLESNHTIKAGHLVTIMEVSRDGTGNTVILT 402  
347 AGDELVKLPVLAAYPPPEFQYKQKALSGRH---SPHALVKKEVTEASTGTTTLALWN 403  
403 PISKEQSHVSLVWVVPVPPGIEKSLISVSDSYQYGTQTTLCTVYAIAPPHHIIHWYQL 462  
404 SAAGLRNRISLELVNVPPIHEKEASSP-SIYSRHSRQALTCTAYGVPLPLSIQHWHP 462  
463 BEEC---ANEPSQNVSTWNPY-CEWRSVEDFQGNKIEVKNQFALIEGKNTVSTLV 518  
463 WTPCKMFAQRSLRRRQQQDLMPQCRDRAVTTQDAVNPIESLDTWTEFVEGKNTVSKLV 522  
519 IQANVSALYKCEAVNKVGRGSRVLSFHYTRGPE---ITLQDPMQPTQESVSLWCTADR 575  
523 IQANVSAMTKCVSNKVGQDERLYFYVYTTIPDGFTIESPSEBELGGQPVLLSCQADS 582  
576 STPENITWYKLGPPQLPIHVGELTPVCKNLDTLWKLNATMFSNSTNDI-----LIM 627  
583 KYEHLRWYRLNLSLTHDAHGNPLLLDCKQV---HLFATPLAASLEEVAPGARHATLSL 638  
628 ELKNASLDQDGYVCLAQDRKTKRHCVVRQVTLVLERVAPITIGNLEQNTTSIGESIYS 687  
639 SIPRVAPEHEGYVCEVQDRRSHDKHCKKYLVSQALEAPRLTQLTDLNVNDSLEMQ 698  
688 CTASGNPPQIMWFKDNETLVEDSGVLVDKGNRNTIRVRKEDGLYTCOACSVLGCAK 747  
699 CLVAGAHAPSIWYKDERLEEKSGVDLADSNOKLSIORVEEDAGRYLCSVCAKGCVN 758  
748 VEAFFITGAQKTNLEIILVGTAVIAMFFWLLLVILRTVKRANGGELKTGYLSIVMD 807  
759 SSASVAVESGEDKGSNEIVLGTGVIAVFFWVLLLLIFCNWRRAPHADIKTYLSIIMD 818  
808 PDELPLDECHERLPYDASKWEPFRDLKLGKPLGRGAFQGVIEADAFGDKTATCTRTAV 867  
819 PGEVPLEEQCEYLSYDASQWEPFRERHLGRVLGYGAFKGVVEASAFGIHKSSCDTAV 878  
868 KMLKEGATHSEHALMSELKILIHGHLLNVNVLGACTKPGCPPLMWIVECKGNLSTY 927  
879 KMLKEGATASEHALMSELKILIHGHLLNVNVLGACTKPGCPPLMWIVECKGNLSTY 938



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.  
OM protein - protein search, using sw model  
Run on: March 14, 2006, 01:40:03 ; Search time 193 Seconds  
(without alignments)  
2558.869 Million cell updates/sec  
Title: US-10-633-742-8  
Perfect score: 6013  
Sequence: 1 MDLSASLVLCGVSLLSGTV.....TTLYKFTYAGIDCSAEAA 1124  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 2443163 seqs, 439378781 residues  
Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries  
Database : A\_Geneseq\_21.\*  
1: Geneseq1980s:\*  
2: Geneseq1990s:\*  
3: Geneseq2000s:\*  
4: Geneseq2001s:\*  
5: Geneseq2002s:\*  
6: Geneseq2003s:\*  
7: Geneseq2003bs:\*  
8: Geneseq2004s:\*  
9: Geneseq2005s:\*  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6013	100.0	1124	2	AAR73953 Human TEK
2	6013	100.0	1124	2	AAY30318 Amino aci
3	6013	100.0	1124	4	AAG65945 Amino aci
4	6013	100.0	1124	5	AAR78550 Human Tie
5	6013	100.0	1124	5	ABB84857 Human PRO
6	6013	100.0	1124	5	ABB95463 Human ang
7	6013	100.0	1124	6	ABU03528 Angiogene
8	6013	100.0	1124	6	ABU07838 Human Tie
9	6013	100.0	1124	7	ADD10371 Human sec
10	6013	100.0	1124	7	ADD11331 Human sec
11	6013	100.0	1124	7	ADD37124 Human sec
12	6013	100.0	1124	7	ADF45095 Human kin
13	6013	100.0	1124	8	ADE41332 Human sec
14	6013	100.0	1124	8	ADH43515 Human PRO
15	6013	100.0	1124	8	ADK82860 Human PRO
16	6013	100.0	1124	8	ADP44301 Human TEK
17	6013	100.0	1124	8	ADR87236 Amino aci
18	6013	100.0	1124	9	ADZ26563 Human tie
19	6013	100.0	1124	9	AEA81524 Human TEK
20	6008	99.9	1124	2	AAR45440 Human orp
21	5598.5	93.1	1123	2	AAR73951 Mouse tie
22	5598.5	93.1	1123	2	AAR67391 Murine ty
23	5575	92.7	1122	3	AAY59048 Mouse tie
24	5556	92.4	1122	2	AAR48627 Protein-t

ALIGNMENTS

RESULT 1  
AAR73953  
ID AAR73953 standard; protein; 1124 AA.  
XX  
AC AAR73953;  
XX  
DT 22-JAN-1996 (first entry)  
XX  
DE Human TEK tyrosine kinase protein.  
XX  
KW tie-2; receptor-tyrosine kinase; DNA primer; cancer; angiogenesis;  
KW vasculogenesis; tek.  
XX  
OS Homo sapiens.  
XX  
PN WO9513387-A1.  
XX  
PD 18-MAY-1995.  
XX  
PF 12-NOV-1994; 94WO-EP003767.  
XX  
PR 12-NOV-1993; 93US-00152552.  
XX  
(PLAC ) MAX PLANCK GBS FOERDERUNG WISSENSCHAFTEN.  
PA  
PI Risau W;  
XX  
DR WPI; 1995-194105/25.  
XX  
PT New tie-2 receptor tyrosine kinase and related nucleic acid - and methods  
PT for detecting tie-2 modulators for treating eg cancer, associated with  
PT angiogenesis and vasculogenesis.  
XX  
XX Disclosure; Page 49; 81pp; English.  
PS This protein is the human homolog of mouse tie-2 receptor tyrosine kinase  
CC  
XX  
SQ Sequence 1124 AA;  
Query Match 100.0%; Score 6013; DB 2; Length 1124;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MDLSASLVLCGVSLLSGTVEGAMDLILNSPLVSDAETSLTCTASGWRPHEPTIGRD 60  
Db 1 MDLSASLVLCGVSLLSGTVEGAMDLILNSPLVSDAETSLTCTASGWRPHEPTIGRD 60



Db 181 AQPQAGVSVARYIGNLTSAFTRLIVRCAQKWGPCNHLCTACMNGVCHEDTGBEC 240  
Qy 241 ICPPGFMGRTCEKACELHTFGRTKERCSCGQCKSYVFCCLPDPYGCSCATCWKGLQCNIE 300  
Db 241 ICPPGFMGRTCEKACELHTFGRTKERCSCGQCKSYVFCCLPDPYGCSCATCWKGLQCNIE 300  
Qy 301 ACHPGFYGDPCKLRCSCNNGECNDRFQGLCSFGWQGLQCEGREGIPRMTPKIVDLPDHE 360  
Db 301 ACHPGFYGDPCKLRCSCNNGECNDRFQGLCSFGWQGLQCEGREGIPRMTPKIVDLPDHE 360  
Qy 361 VNSGKFNPKCKASGPNLPNEBMTLVKPDGTVLHPKDFNHTDHFSAIPIHRIILPPDSG 420  
Db 361 VNSGKFNPKCKASGPNLPNEBMTLVKPDGTVLHPKDFNHTDHFSAIPIHRIILPPDSG 420  
Qy 421 VVWCSVNTVAGVVEKPFNIVSVKLPRKLNAPNVIDTGHNFVAVINISSEBYPFGDPIKSKK 480  
Db 421 VVWCSVNTVAGVVEKPFNIVSVKLPRKLNAPNVIDTGHNFVAVINISSEBYPFGDPIKSKK 480  
Qy 481 LLYKPNHYEAWQHIVQTNIEIVTLNLEPRTEYELCVQLVRRGEGEGHPGVRRFTTAS 540  
Db 481 LLYKPNHYEAWQHIVQTNIEIVTLNLEPRTEYELCVQLVRRGEGEGHPGVRRFTTAS 540  
Qy 541 IGLPPRGNLNLPKSGOTTLNLTWQPIFPSSDDFVVEVERRSVQKSDQONIKVPGNLTSV 600  
Db 541 IGLPPRGNLNLPKSGOTTLNLTWQPIFPSSDDFVVEVERRSVQKSDQONIKVPGNLTSV 600  
Qy 601 LNNLHPREQYVVRVAVNTKAQWSEDLTAWTSLDILPPQENIKISNITHSSAVISMT 660  
Db 601 LNNLHPREQYVVRVAVNTKAQWSEDLTAWTSLDILPPQENIKISNITHSSAVISMT 660  
Qy 661 ILDGYSISITIRYKVGKNEQHVVDVKIKNATIIQYQLKLEPETAYQVDIFAENNIGS 720  
Db 661 ILDGYSISITIRYKVGKNEQHVVDVKIKNATIIQYQLKLEPETAYQVDIFAENNIGS 720  
Qy 721 SNPAFSEHVLTPESQAPADLGGKWLAIIGSAGTCLTVLLAFLIILQKRVAVRR 780  
Db 721 SNPAFSEHVLTPESQAPADLGGKWLAIIGSAGTCLTVLLAFLIILQKRVAVRR 780  
Qy 781 MAQAFQNVREPAVQFNSGTLALNRKVNPNPDTIYVLDNDIKFDQVIGEGNGQVLK 840  
Db 781 MAQAFQNVREPAVQFNSGTLALNRKVNPNPDTIYVLDNDIKFDQVIGEGNGQVLK 840  
Qy 841 ARIKQDGLRMDAAIKRMKEYASKDHRDPAGBLEVLCKLGHHPNIINLLGACEHRYLYL 900  
Db 841 ARIKQDGLRMDAAIKRMKEYASKDHRDPAGBLEVLCKLGHHPNIINLLGACEHRYLYL 900  
Qy 901 ALEYAPHGNLDPKRSVLETDPAFIANSFASITLSSQALLHFAADVARGMDYLSQKQF 960  
Db 901 ALEYAPHGNLDPKRSVLETDPAFIANSFASITLSSQALLHFAADVARGMDYLSQKQF 960  
Qy 961 IHRDLAARNILGVENVYAKIADFGLSRGQEVYVKKTMGRLPVRWMAIESLNTSVYTTNSD 1020  
Db 961 IHRDLAARNILGVENVYAKIADFGLSRGQEVYVKKTMGRLPVRWMAIESLNTSVYTTNSD 1020  
Qy 1021 VWSYGVLLWEIVSLGTPYCGMTCAELYEKLPGQYRLEKPLNCDDBEVYDLMOQCRKPY 1080  
Db 1021 VWSYGVLLWEIVSLGTPYCGMTCAELYEKLPGQYRLEKPLNCDDBEVYDLMOQCRKPY 1080  
Qy 1081 ERPSFAQILVSLNRMLEERKTVNTTLYEKFTYVAGIDCSAEBA 1124  
Db 1081 ERPSFAQILVSLNRMLEERKTVNTTLYEKFTYVAGIDCSAEBA 1124

RESULT 3

ID AAG65945 standard; protein; 1124 AA.

AC AAG65945;

XX 11-FEB-2002 (first entry)

DT Amino acid sequence of human Tie-2.

XX

Tie-2; catalytic domain; receptor tyrosine kinase; crystalline; human; cytosolic; vasotrophic; antianemic; antihypertensive; nephrotropic; opthalmologic; hepatotropic; antithyroid; antiinflammatory; antitumor; gastrointestinal; antirheumatic; osteopathic; antiarthritic; hemostatic; antipsoriatic; dermatological; immunosuppressive; antibacterial.  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Domain 802. .1124  
FT /note= "catalytic domain; specifically claimed fragment"  
XX  
XX WO200172778-A2.  
PD 04-OCT-2001.  
XX  
XX 20-MAR-2001; 2001WO-US008853.  
XX 29-MAR-2000; 2000US-0192920P.  
XX (BADI ) BASF AG.  
XX Bump NJ, Arnold LD, Dixon RW, Hoeffken HW, Allen K;  
XX Bellamacina C;  
XX WPI; 2001-648437/74.  
XX  
XX Crystalline polypeptide useful for identifying inhibitors of a Tie-2 protein as well as determining the three dimensional structure of a polypeptide comprising the catalytic domain of a Tie-2 polypeptide.  
XX  
XX Claim 6; Fig 1; 242pp; English.  
XX  
XX The invention relates to a crystalline polypeptide, comprising the catalytic domain of a receptor tyrosine kinase Tie-2 protein. The crystalline forms are useful for identifying inhibitors of a Tie-2 protein as well as determining the three dimensional structure of a the catalytic domain of a Tie-2 polypeptide. A Tie-2 inhibitor may be used to treat a Tie-2 dependent condition in a patient (especially a human), where the condition is characterized by excessive vascular proliferation e.g. a hyperproliferative disorder, cancer (e.g. sarcoma, osteoma, melanoma, lymphoma, and leukemia), a cardiovascular condition (e.g. atherosclerosis, ischemia, anemia, and vascular leakage disorders), an ocular condition (myopia, chronic retinal detachment, conjunctivitis, retinopathy, and macular degeneration), von Hippel Lindau disease, pemphigoid, psoriasis, Paget's disease, polycystic kidney disease, fibrosis, sarcoidosis, cirrhosis, thyroiditis, Osler-Weber-Rendu disease, chronic inflammation, synovitis, inflammatory bowel disease, Crohn's disease, rheumatoid arthritis, osteoarthritis, psoriatic arthritis, and ulcer or sepsis, especially where the disorder involves aberrant endothelial-periendothelial interactions. The Tie-2 inhibitor may be used to decrease fertility, and promote angiogenesis or vasculogenesis (in combination with a pro-angiogenic growth factor). The present sequence represents the human Tie-2 protein sequence

XX Sequence 1124 AA;

Query Match 100.0%; Score 6013; DB 4; Length 1124;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDLSASLVLCVSGVLLSGTVEGAMDILLINSPLVSDAETSITCTASGWRPHEPTIGRD 60

Db 1 MDLSASLVLCVSGVLLSGTVEGAMDILLINSPLVSDAETSITCTASGWRPHEPTIGRD 60

Qy 61 FEALNNQHPDPLEVTQDVTREWAKVWVKREKASKINGAYFCGEVRGGEAIRRTMKMQ 120

Db 61 FEALNNQHPDPLEVTQDVTREWAKVWVKREKASKINGAYFCGEVRGGEAIRRTMKMQ 120

Qy 121 QASFLPATLTMTVDKGDNNVTSFKKVLKEEDAVYKNGSFTHSVPRHEVPDILEVHLPH 180

Db 121 QASFLPATLTMTVDKGDNNVTSFKKVLKEEDAVYKNGSFTHSVPRHEVPDILEVHLPH 180

QY 181 AQPQAGVYSARYIGNLFTSAFTRLIVRCEBAQKWGPCNHLCTACMNGVCHEDTGBC 240  
Db 181 AQPQAGVYSARYIGNLFTSAFTRLIVRCEBAQKWGPCNHLCTACMNGVCHEDTGBC 240  
QY 241 ICPPGFMGRTCEKACELHHTFGRTKERCSCGQCKSYVFCCLPDPYGCSCATCKWGLQCN 300  
Db 241 ICPPGFMGRTCEKACELHHTFGRTKERCSCGQCKSYVFCCLPDPYGCSCATCKWGLQCN 300  
QY 301 ACHPGFYGDPCKLRCSCNNGECDFRQGLCSFGHQGLQCEGIPRMTPKIVLDPDHE 360  
Db 301 ACHPGFYGDPCKLRCSCNNGECDFRQGLCSFGHQGLQCEGIPRMTPKIVLDPDHE 360  
QY 361 VNSGKFNPICKASGWLPTNEBMTLVKPDGTVLHPKDFNHTDHFSAIPIHRIILPPDSG 420  
Db 361 VNSGKFNPICKASGWLPTNEBMTLVKPDGTVLHPKDFNHTDHFSAIPIHRIILPPDSG 420  
QY 421 VVWCSVNTVAGMVEKPFNISKVLKPLNAPNVIDTGHNFVAINISSPEYFGDGIKSKK 480  
Db 421 VVWCSVNTVAGMVEKPFNISKVLKPLNAPNVIDTGHNFVAINISSPEYFGDGIKSKK 480  
QY 481 LLYKPNHYEAWOHIQVNTNEIYTLNLEPRTYEYELCVQLVRRGEGEGHPGVRFTTAS 540  
Db 481 LLYKPNHYEAWOHIQVNTNEIYTLNLEPRTYEYELCVQLVRRGEGEGHPGVRFTTAS 540  
QY 541 IGLPPRGLNLKPKSOTTLNLTWQPIFPSSDDFVVEVERRSVQKSDQONIKVPGNLTSV 600  
Db 541 IGLPPRGLNLKPKSOTTLNLTWQPIFPSSDDFVVEVERRSVQKSDQONIKVPGNLTSV 600  
QY 601 LNNLHPREQYVVRARVNTKAQGESEDLTAWTSDILPPQENIKISNITHSSAVISWT 660  
Db 601 LNNLHPREQYVVRARVNTKAQGESEDLTAWTSDILPPQENIKISNITHSSAVISWT 660  
QY 661 ILDGYSISSITIRYKVGQKNEHDVVKIKNATIIQYQLKLEPETAYQVDIFAENNTGS 720  
Db 661 ILDGYSISSITIRYKVGQKNEHDVVKIKNATIIQYQLKLEPETAYQVDIFAENNTGS 720  
QY 721 SNPAFSELVTLPEQAPADLGGKMLIAILGSAGMTCLTVLLAFLIILQLKRANVQR 780  
Db 721 SNPAFSELVTLPEQAPADLGGKMLIAILGSAGMTCLTVLLAFLIILQLKRANVQR 780  
QY 781 MAQAFQNVREEPAPVFNQSGTTLALNRKVNKNPDPPTIYPVLDWMDIKFQDVIGEGNFGVLK 840  
Db 781 MAQAFQNVREEPAPVFNQSGTTLALNRKVNKNPDPPTIYPVLDWMDIKFQDVIGEGNFGVLK 840  
QY 841 ARIKXDLGRMDAAIKRMKEYASKDDHDPAGELEVLCKLGHHPNIIINLLGACEHRYLYL 900  
Db 841 ARIKXDLGRMDAAIKRMKEYASKDDHDPAGELEVLCKLGHHPNIIINLLGACEHRYLYL 900  
QY 901 ALEYAPHGNLLDPLKRSRVLETDPAFANSTASTLSSQQLHFAADVARGMDYLSQKF 960  
Db 901 ALEYAPHGNLLDPLKRSRVLETDPAFANSTASTLSSQQLHFAADVARGMDYLSQKF 960  
QY 961 IHRDLAARNILVGENYAKIADFGLSRGQEVVYKKTGMRLPVRWMAIBSLNYSVYTTNSD 1020  
Db 961 IHRDLAARNILVGENYAKIADFGLSRGQEVVYKKTGMRLPVRWMAIBSLNYSVYTTNSD 1020  
QY 1021 VMSYGVLLWEIVISLGGTPCYGTCMAELYEKLPQGRYLEKPLNCDDDEVYDLMRQCRKPY 1080  
Db 1021 VMSYGVLLWEIVISLGGTPCYGTCMAELYEKLPQGRYLEKPLNCDDDEVYDLMRQCRKPY 1080  
QY 1081 ERPSFAQILVSNLRMLEERKTVNTTLYEKFTYAGIDCSAEAA 1124  
Db 1081 ERPSFAQILVSNLRMLEERKTVNTTLYEKFTYAGIDCSAEAA 1124

RESULT 4

AAU78550

XX AAU78550 standard; protein; 1124 AA.

AC AAU78550;

XX

DT 18-JUN-2002 (first entry)

XX

DE Human Tie2 receptor tyrosine kinase protein.  
XX Tie2; Tie2K; human; crystal structure; vascular dysmorphogenesis;  
KW protein co-ordinate data; receptor tyrosine kinase; vascular development.  
XX Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain 808..1124  
FT Binding-site /note= "Tie2K receptor kinase domain"  
FT /note= "ATP binding site"  
FT Region 897  
FT /note= "Phosphorylation site"  
FT Region 964  
FT /note= "Catalytic aspartate residue"  
FT Region 982..1008  
FT /note= "Activation loop"  
FT Region 1119  
FT /note= "Phosphorylation site"  
XX WO200220734-A2.  
FN 14-MAR-2002.  
PD  
XX 06-SEP-2001; 2001WO-US027486.  
PF  
XX 08-SEP-2000; 2000US-0231398P.  
PR (GLAX ) GLAXO GROUP LTD.  
XX  
PI Davis RG, Ellis BP, Hassell AM, Holmes WD, Shewchuk LM;  
XX WPI; 2002-292263/33.  
DR N-PSDB; ABK12246.  
XX  
PT Composition for the design or screening of a cytoplasmic Tie2 receptor  
PT tyrosine kinase domain modulator, comprises a monoclinic or orthorhombic  
PT crystalline form of a cytoplasmic Tie2 receptor tyrosine kinase domain  
PT polypeptide.  
XX  
PS Disclosure; Page 565-570; 581pp; English.  
XX  
CC This invention relates to the 3 dimensional crystal structure of the  
CC human Tie2 receptor tyrosine kinase domain polypeptide in a monoclinic,  
CC or orthorhombic crystalline form. Tie2 is an endothelial-specific  
CC receptor tyrosine kinase thought to be involved in vascular development.  
CC Naturally occurring R849W and Y897S mutations in the Tie2 protein have  
CC been identified in human s and have been shown to segregate with the  
CC autosomal dominant condition vascular dysmorphogenesis. The invention  
CC also comprises methods for designing modulators of the biological  
CC activity of the cytoplasmic Tie2 receptor tyrosine kinase domain. The  
CC Tie2 receptor tyrosine kinase domain is used to design or screen for a  
CC modulator of the kinase by rational drug design, using computer models.  
CC The present sequence represents the human Tie2 protein used to create the  
CC Tie2K tyrosine kinase domain of the invention  
XX  
SQ Sequence 1124 AA;  
Query Match 100.0%; Score 6013; DB 5; Length 1124;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDSLASIVLCGVSLLSGTVEGAMDILILNSILPLVSDAETSLTCIASCGRPREPTIGRD 60  
Db 1 MDSLASIVLCGVSLLSGTVEGAMDILILNSILPLVSDAETSLTCIASCGRPREPTIGRD 60  
QY 61 FEALMNQHQDPLEVTQDVTREWAKVWVKREKASKINGAYFCEGVRGGEAIRITWKMQRQ 120  
Db 61 FEALMNQHQDPLEVTQDVTREWAKVWVKREKASKINGAYFCEGVRGGEAIRITWKMQRQ 120  
QY 121 QASFLPATLTWTVDKGDNNVISFKVLKIBEDAVTYKNGSFTHSVPRHEVPDILEVHLPH 180  
XX



Db 121 QASFLPATLTMTVDKDNVNI SPKVL I KEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
Qy 181 AQPQDAGVYSARYIGGNLFSAFTRLI VRRCEAQKMGPECNHLCTACMNGVCHEDTGC 240  
Db 181 AQPQDAGVYSARYIGGNLFSAFTRLI VRRCEAQKMGPECNHLCTACMNGVCHEDTGC 240  
Qy 241 ICPPGFMGTCEKACBLHFTGRTCKRCGQBECKSYVFLPDPYCSCATGWKGLQCN 300  
Db 241 ICPPGFMGTCEKACBLHFTGRTCKRCGQBECKSYVFLPDPYCSCATGWKGLQCN 300  
Qy 301 ACHPGYGPDPCKLRCSCNNGEMCDRFQGLCSFGWGLQCEREGIPRMTPKVLDLPH 360  
Db 301 ACHPGYGPDPCKLRCSCNNGEMCDRFQGLCSFGWGLQCEREGIPRMTPKVLDLPH 360  
Qy 361 VNSGKNPICKASGWLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAIFTIHRILPPDSG 420  
Db 361 VNSGKNPICKASGWLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAIFTIHRILPPDSG 420  
Qy 421 VVWCSVNTVAGWVEKPFNISVKVLPKPLNAPNVIDTGHNPFAVINISSEPYFGDGPISK 480  
Db 421 VVWCSVNTVAGWVEKPFNISVKVLPKPLNAPNVIDTGHNPFAVINISSEPYFGDGPISK 480  
Qy 481 LLYKPNHYEAMQHIOVTNEI VTLNLEPRTEVELCVQLVRRGEGEGHPVRRFTTAS 540  
Db 481 LLYKPNHYEAMQHIOVTNEI VTLNLEPRTEVELCVQLVRRGEGEGHPVRRFTTAS 540  
Qy 541 IGLPPRGLNLAPKSTTLNLWTQPIFPSSSEDDFYVEVERRSVQKSDQDNIPVGNLTSV 600  
Db 541 IGLPPRGLNLAPKSTTLNLWTQPIFPSSSEDDFYVEVERRSVQKSDQDNIPVGNLTSV 600  
Qy 601 LNLNLHPRQYVVRVNTKAGWSEDLTAMTSLDILPPQENIKISNITHSSAVISMT 660  
Db 601 LNLNLHPRQYVVRVNTKAGWSEDLTAMTSLDILPPQENIKISNITHSSAVISMT 660  
Qy 661 ILDGYSISITIRYKQGNEDQHDVVKIKNATIIQYQLKGLEPETAYQVDIIFAENNIGS 720  
Db 661 ILDGYSISITIRYKQGNEDQHDVVKIKNATIIQYQLKGLEPETAYQVDIIFAENNIGS 720  
Qy 721 SNPAFSHELVTLPESQAPADLGGKMLLAILGSAGMTCLTVLLAPLIILQKRANVQR 780  
Db 721 SNPAFSHELVTLPESQAPADLGGKMLLAILGSAGMTCLTVLLAPLIILQKRANVQR 780  
Qy 781 MAQAFQNVREPAVQNSGTLALNRKKNPDPPTIYPVLDWNDIKFQDVIGSGNFGVLK 840  
Db 781 MAQAFQNVREPAVQNSGTLALNRKKNPDPPTIYPVLDWNDIKFQDVIGSGNFGVLK 840  
Qy 841 ARIKXGDLRMDAAIKMKKEYASKDHRDFAGELEVLCKLGHHPNIIINLIGACEHGRYL 900  
Db 841 ARIKXGDLRMDAAIKMKKEYASKDHRDFAGELEVLCKLGHHPNIIINLIGACEHGRYL 900  
Qy 901 ALEYAPHGNLLDFLRKSRVLETPAPAIANSTASTLSSQQLHFAADVARGMDYLSQKF 960  
Db 901 ALEYAPHGNLLDFLRKSRVLETPAPAIANSTASTLSSQQLHFAADVARGMDYLSQKF 960  
Qy 961 IHRDLAARNILGENTYAKIADFGLSRGQEVYKKTGMRLPVRRWMAIESLNTSVYTTNSD 1020  
Db 961 IHRDLAARNILGENTYAKIADFGLSRGQEVYKKTGMRLPVRRWMAIESLNTSVYTTNSD 1020  
Qy 1021 VNSYGVLLWEIYSLGTPYCGMTCAELYEKLPQGYLEKPLNCDDBVYDLMRQCWREKY 1080  
Db 1021 VNSYGVLLWEIYSLGTPYCGMTCAELYEKLPQGYLEKPLNCDDBVYDLMRQCWREKY 1080  
Qy 1081 ERPSFAQILVSLNRLMEERKTYNTTLYEKFTYAGIDCSAEAA 1124  
Db 1081 ERPSFAQILVSLNRLMEERKTYNTTLYEKFTYAGIDCSAEAA 1124

RESULT 5

ID ABB84857

XX ABB84857 standard; protein; 1124 AA.

AC ABB84857;

XX

DT 16-MAY-2002 (first entry)  
XX Human PRO734 protein sequence SEQ ID NO:82.  
XX Human; angiogenesis; cardiant; cytostatic; antiangiogenic; hypotensive;  
XX vulnerary; antiarteriosclerotic; PRO agonist; PRO antagonist; trauma;  
XX gene therapy; cardiovascular disorder; endothelial disorder; cancer;  
XX angiogenic disorder; cardiac hypertrophy; atherosclerosis; hypertension;  
XX age-related macular degeneration; arterial restenosis; angina;  
XX rheumatoid arthritis; myocardial infarction; thrombophlebitis;  
XX lymphangitis; tumour angiogenesis; breast carcinoma; liver carcinoma;  
XX wound healing; chromosome mapping; gene mapping.  
OS Homo sapiens.  
XX WO200200690-A2.  
XX 03-JAN-2002.  
XX 20-JUN-2001; 2001WO-US019692.  
XX 23-JUN-2000; 2000US-0213637P.  
XX 20-JUL-2000; 2000US-0219556P.  
XX 25-JUL-2000; 2000US-0220624P.  
XX 25-JUL-2000; 2000US-0220664P.  
XX 28-JUL-2000; 2000WO-US020710.  
XX 02-AUG-2000; 2000US-0222695P.  
XX 17-AUG-2000; 2000US-00643657.  
XX 23-AUG-2000; 2000WO-US023522.  
XX 24-AUG-2000; 2000WO-US023328.  
XX 07-SEP-2000; 2000US-0230978P.  
XX 18-SEP-2000; 2000US-02066410.  
XX 18-SEP-2000; 2000US-00665350.  
XX 24-OCT-2000; 2000US-0242922P.  
XX 08-NOV-2000; 2000US-00709238.  
XX 08-NOV-2000; 2000WO-US030932.  
XX 10-NOV-2000; 2000WO-US030873.  
XX 01-DEC-2000; 2000WO-US032678.  
XX 20-DEC-2000; 2000US-00747259.  
XX 20-DEC-2000; 2000WO-US034956.  
XX 22-JAN-2001; 2001US-00767609.  
XX 28-FEB-2001; 2001US-00796498.  
XX 28-FEB-2001; 2001WO-US006520.  
XX 01-MAR-2001; 2001WO-US006666.  
XX 09-MAR-2001; 2001US-00802706.  
XX 14-MAR-2001; 2001US-00808689.  
XX 22-MAR-2001; 2001US-00816744.  
XX 05-APR-2001; 2001US-00828366.  
XX 10-MAY-2001; 2001US-00854208.  
XX 25-MAY-2001; 2001US-00866028.  
XX 25-MAY-2001; 2001US-00866034.  
XX 25-MAY-2001; 2001WO-US017092.  
XX 30-MAY-2001; 2001US-00870574.  
XX 30-MAY-2001; 2001WO-US017443.  
XX 01-JUN-2001; 2001WO-US017800.  
XX (GETH ) GENENTECH INC.  
XX Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A;  
PI Godowski PJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Paoni NF;  
PI Stephan JF, Watanabe CK, Williams PM, Wood WI, Ye W;  
XX WPI; 2002-090516/12.  
XX N-PSDB; ABL88112.  
XX One hundred and eighty seven nucleic acids encoding PRO polypeptides,  
XX useful in diagnosis and treatment of cardiovascular (e.g. myocardial  
XX infarction), endothelial or angiogenic disorders in a mammal.  
XX Claim 11; Fig 82; 565pp; English.  
XX ABL88072 to ABL88258 encode the PRO proteins given in ABB84817 to  
XX

CC ABB85003. The PRO proteins and polynucleotides have cardiant, cytostatic,  
CC antiangiogenic, hypotensive, vulnary and antiarteriosclerotic  
CC activities, and can be used in gene therapy. The PRO polynucleotides,  
CC proteins, agonists and antagonists are useful for treating or diagnosing  
CC a cardiovascular, endothelial or angiogenic disorder in a mammal, e.g.  
CC cardiac hypertrophy, trauma, cancer, age-related macular degeneration,  
CC atherosclerosis, hypertension, arterial restenosis, rheumatoid arthritis,  
CC angina, myocardial infarctions, thrombophlebitis, lymphangitis, tumour  
CC angiogenesis (such as breast carcinoma and liver carcinoma) and wound  
CC healing. The PRO polynucleotides have applications in molecular biology,  
CC including use as hybridisation probes, and in chromosome and gene  
CC mapping. ABL88259 to ABL88267 represent primers and probes used in the  
CC exemplification of the present invention  
XX  
SQ Sequence 1124 AA;

Query Match		100.0%;	Score 6013;	DB 5;	Length 1124;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1124;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MDSLASVLCGVSLLSGTVEGAMDILILINSIPLVSDAETSLTCTASGWRPHEPTIGRD	60		
DB	1	MDSLASVLCGVSLLSGTVEGAMDILILINSIPLVSDAETSLTCTASGWRPHEPTIGRD	60		
QY	61	FEALMNQHDPLEVTQDVTREWAKVWVKREKASKINGAYFCGEGRVGEAIRIRTMKVRQ	120		
DB	61	FEALMNQHDPLEVTQDVTREWAKVWVKREKASKINGAYFCGEGRVGEAIRIRTMKVRQ	120		
QY	121	QASFLPATLTMTVDKGDVNIISFKVLIKEEDAVYKNGSPFIHSVPRHEVPDILEVHLPH	180		
DB	121	QASFLPATLTMTVDKGDVNIISFKVLIKEEDAVYKNGSPFIHSVPRHEVPDILEVHLPH	180		
QY	181	AQPDAGVYSARYIGNLFTSAFTRLIVRCEAQKWGPCNHLCTACMNGVCHEDTGEC	240		
DB	181	AQPDAGVYSARYIGNLFTSAFTRLIVRCEAQKWGPCNHLCTACMNGVCHEDTGEC	240		
QY	241	ICPPFGMGRTCCKACELHFTGRTCKERCSCGQCKSYVFCCLPDPYGCSCATCWKGLQCN	300		
DB	241	ICPPFGMGRTCCKACELHFTGRTCKERCSCGQCKSYVFCCLPDPYGCSCATCWKGLQCN	300		
QY	301	ACHPGFYGPDKLRCSCNNGECMDRFPQGLCSFGWQGLQCEGREGIPRMTPKIVLDPDHE	360		
DB	301	ACHPGFYGPDKLRCSCNNGECMDRFPQGLCSFGWQGLQCEGREGIPRMTPKIVLDPDHE	360		
QY	361	VNSGFNPKCKASGWPLFPNEEMTLVKPDGTVLHPKDFNHTDHFVAIETHRIILPPDSG	420		
DB	361	VNSGFNPKCKASGWPLFPNEEMTLVKPDGTVLHPKDFNHTDHFVAIETHRIILPPDSG	420		
QY	421	VWVCSVNTVAGWVEKPFNLSVKVLPKPLNAPNVIDTGHNFVAINISSEPYFGDGPIKSKK	480		
DB	421	VWVCSVNTVAGWVEKPFNLSVKVLPKPLNAPNVIDTGHNFVAINISSEPYFGDGPIKSKK	480		
QY	481	LLYKPNHYEAQHIOVTNEIVTLNLYLEPRTEYELCVQLVRGEGEGHPGVRPFTTAS	540		
DB	481	LLYKPNHYEAQHIOVTNEIVTLNLYLEPRTEYELCVQLVRGEGEGHPGVRPFTTAS	540		
QY	541	IGLPPRGNLAPKSGTTLNLTWQPIFSSDDDFVVEVERRSVQSDQONIKVPGNLTSV	600		
DB	541	IGLPPRGNLAPKSGTTLNLTWQPIFSSDDDFVVEVERRSVQSDQONIKVPGNLTSV	600		
QY	601	LNLNHPRQYVVRVNTVKAQWSEDTATWTLSDILPPQENIKISNITHSSAVISWT	660		
DB	601	LNLNHPRQYVVRVNTVKAQWSEDTATWTLSDILPPQENIKISNITHSSAVISWT	660		
QY	661	ILDGYSISSITIRYKQGNEDQHDVVKIKNATIIQYQLKLEPETAYQVDIPFAENNIGS	720		
DB	661	ILDGYSISSITIRYKQGNEDQHDVVKIKNATIIQYQLKLEPETAYQVDIPFAENNIGS	720		
QY	721	SNAPASHELVTLIPESQAPADLGGKMLLIALIGSAGMTCLTVLLAFLIILQKRVORR	780		
DB	721	SNAPASHELVTLIPESQAPADLGGKMLLIALIGSAGMTCLTVLLAFLIILQKRVORR	780		
QY	781	MAQAFQNVREBPVNSGTLALNRKVNKNPDPTIYPVLDWMDIKFQDVIGEGNFQVLK	840		

Db	781	MAQAFQNVREBPVNSGTLALNRKVNKNPDPTIYPVLDWMDIKFQDVIGEGNFQVLK	840	
Qy	841	ARIKKDGLRMDAAIKEMKEYASKDHRDPAGELEVLCKLGHHPNINILLGACEHRGYLYL	900	
Db	841	ARIKKDGLRMDAAIKEMKEYASKDHRDPAGELEVLCKLGHHPNINILLGACEHRGYLYL	900	
Qy	901	AIETYAPHGNLLDFLRKSRVLETFDPAFAIANSTASTLSSQQLHFAADVARGMDYLSQKF	960	
Db	901	AIETYAPHGNLLDFLRKSRVLETFDPAFAIANSTASTLSSQQLHFAADVARGMDYLSQKF	960	
Qy	961	IHRDLAARNILVGENYVAKIADFGLSRGQEVVYKVTMGRLPVRWMAIESLNSVYTTNSD	1020	
Db	961	IHRDLAARNILVGENYVAKIADFGLSRGQEVVYKVTMGRLPVRWMAIESLNSVYTTNSD	1020	
Qy	1021	VMSYGYLLWEIVSLGTPYCGMTCAELYEKLPGQYRLEKPLNCDDEYDLMRCQWREKPY	1080	
Db	1021	VMSYGYLLWEIVSLGTPYCGMTCAELYEKLPGQYRLEKPLNCDDEYDLMRCQWREKPY	1080	
Qy	1081	ERPSFAQILVSLNRMLEERKTYVNTTLYEKFTYAGIDCSAEAAA	1124	
Db	1081	ERPSFAQILVSLNRMLEERKTYVNTTLYEKFTYAGIDCSAEAAA	1124	
RESULT 6				
ABB95463				
XX	ID	ABB95463 standard; protein; 1124 AA.		
XX	AC	ABB95463;		
XX	DT	19-JUL-2002 (first entry)		
XX	DE	Human angiogenesis related protein PRO734 SEQ ID NO: 82.		
XX	KW	Human; angiogenesis; PRO protein; cardiovascularisation; wound; cancer; atherosclerosis; cardiac hypertrophy; gene therapy; endothelial disorder; cardiant; cytostatic; antiangiogenic; hypotensive; vulnary;		
XX	KW	antiarteriosclerotic.		
XX	OS	Homo sapiens.		
XX	OS			
XX	PN	WO200208284-A2.		
XX	PD	31-JAN-2002.		
XX	PF	09-JUL-2001; 2001WO-US021735.		
XX	PR	20-JUL-2000; 2000US-0219556P.		
XX	PR	25-JUL-2000; 2000US-0220624P.		
XX	PR	25-JUL-2000; 2000US-0220664P.		
XX	PR	28-JUL-2000; 2000WO-US020710.		
XX	PR	02-AUG-2000; 2000US-0222695P.		
XX	PR	17-AUG-2000; 2000US-00643657.		
XX	PR	23-AUG-2000; 2000WO-US023522.		
XX	PR	24-AUG-2000; 2000WO-US023328.		
XX	PR	07-SEP-2000; 2000US-0230978P.		
XX	PR	18-SEP-2000; 2000US-00664610.		
XX	PR	24-SEP-2000; 2000US-00655350.		
XX	PR	24-OCT-2000; 2000US-0242922P.		
XX	PR	08-NOV-2000; 2000US-00709238.		
XX	PR	08-NOV-2000; 2000WO-US030952.		
XX	PR	10-NOV-2000; 2000WO-US030873.		
XX	PR	01-DEC-2000; 2000WO-US032678.		
XX	PR	20-DEC-2000; 2000US-00747259.		
XX	PR	20-DEC-2000; 2000WO-US034956.		
XX	PR	22-JAN-2001; 2001US-00767609.		
XX	PR	28-FEB-2001; 2001US-00796498.		
XX	PR	28-FEB-2001; 2001WO-US006520.		
XX	PR	01-MAR-2001; 2001WO-US006666.		
XX	PR	09-MAR-2001; 2001US-00802706.		
XX	PR	14-MAR-2001; 2001US-00808689.		
XX	PR	22-MAR-2001; 2001US-00816744.		
XX	PR	05-APR-2001; 2001US-00828366.		

PR 10-MAY-2001; 2001US-00854208.  
PR 25-MAY-2001; 2001US-00854280.  
PR 25-MAY-2001; 2001US-00866028.  
PR 25-MAY-2001; 2001US-00866034.  
PR 30-MAY-2001; 2001WO-US017092.  
PR 30-MAY-2001; 2001US-00870574.  
PR 01-JUN-2001; 2001WO-US017443.  
PR 20-JUN-2001; 2001WO-US019692.  
XX  
PA (GETH /) GENENTECH INC.  
PA (BAKE /) BAKER K P.  
PA (FERK /) FERRARA N.  
PA (GERB /) GERBER H.  
PA (GERR /) GERRITSEN M E.  
PA (GODD /) GODDARD A.  
PA (GODO /) GODOWSKI P J.  
PA (GURN /) GURNEY A L.  
PA (HILL /) HILLMAN K J.  
PA (MARS /) MARSTERS S A.  
PA (PANJ /) PAN J.  
PA (PAON /) PAONI N F.  
PA (STEP /) STEPHAN J F.  
PA (WATA /) WATANABE C K.  
PA (WILL /) WILLIAMS P M.  
PA (WOOD /) WOOD W I.  
XX  
PI Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A;  
PI Godowski RJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Paoni NF;  
PI Stephan JF, Watanabe CK, Williams PM, Wood WI, Ye W;  
XX  
DR WPI; 2002-171999/22.  
DR N-PSDB; ABL95601.  
XX  
PT One hundred and eighty seven nucleic acids encoding PRO polypeptides,  
PT useful in diagnosis and treatment of cardiovascular (e.g. myocardial  
PT infarction), endothelial or angiogenic disorders in a mammal.  
XX  
PS Claim 11; Fig 82; 567pp; English.  
XX  
CC The present invention provides the protein and coding sequences of human  
CC PRO proteins. These are useful for treating or diagnosing a  
CC cardiovascular, endothelial or angiogenic disorder, including cardiac  
CC hypertrophy, trauma, cancer, age-related macular degeneration,  
CC atherosclerosis, hypertension, arterial restenosis, rheumatoid arthritis,  
CC angina, myocardial infarctions, thrombophlebitis, lymphangitis, tumour  
CC angiogenesis (such as breast carcinoma and liver carcinoma) and wound  
CC healing. The present sequence is a PRO protein of the invention  
XX  
SQ Sequence 1124 AA;  
Query Match 100.0%; Score 6013; DB 5; Length 1124;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDSLASLVLCGVSLLSGTVEGAMDILINSLPLVSDAETSLTCTIASGWRPHEPTIGRD 60  
DB 1 MDSLASLVLCGVSLLSGTVEGAMDILINSLPLVSDAETSLTCTIASGWRPHEPTIGRD 60  
QY 61 FEALMNHQDPLEVTQDVTREWAKVWVREKASKINGAYFCEGRVGRGEAIRITWKRQ 120  
DB 61 FEALMNHQDPLEVTQDVTREWAKVWVREKASKINGAYFCEGRVGRGEAIRITWKRQ 120  
QY 121 QASFLPATLTMTVDKGNVNISPKVLIKEEDAVIYKNGSFTHSVPRHVPDILEVHLPH 180  
DB 121 QASFLPATLTMTVDKGNVNISPKVLIKEEDAVIYKNGSFTHSVPRHVPDILEVHLPH 180  
QY 181 AQPQDAGVSARYIGGNLFTSAFTRLIVRRCCEAQKMGPECNHLCTACMNGVCHEDTGEC 240  
DB 181 AQPQDAGVSARYIGGNLFTSAFTRLIVRRCCEAQKMGPECNHLCTACMNGVCHEDTGEC 240  
QY 241 ICPGPMGRTCBKACELHTFGRTCKERCSCGQCKSVYFCLPDPYGCSCATGWKGLQCNE 300

DB 241 ICPGPMGRTCBKACELHTFGRTCKERCSCGQCKSVYFCLPDPYGCSCATGWKGLQCNE 300  
QY 301 ACHPGFYGPDCKLRCSCNNGEMCDRFGQCLCSFGWQGLQCEGREGIPRMTPKIVDLDPDHIE 360  
DB 301 ACHPGFYGPDCKLRCSCNNGEMCDRFGQCLCSFGWQGLQCEGREGIPRMTPKIVDLDPDHIE 360  
QY 361 VNSGKFNPICKASGWPPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAIETIHRILPPDSG 420  
DB 361 VNSGKFNPICKASGWPPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAIETIHRILPPDSG 420  
QY 421 VVWCVSVTVAGWVEKPFNISVKVLPKLNAPNVIDTGHNFPAVINISSEPVFGDGPISKX 480  
DB 421 VVWCVSVTVAGWVEKPFNISVKVLPKLNAPNVIDTGHNFPAVINISSEPVFGDGPISKX 480  
QY 481 LLYKPNHYEAWQHIQVTNEIVTLNLYLEPRTYELCVQLVRRGEGEGHPGVPVRRFTTAS 540  
DB 481 LLYKPNHYEAWQHIQVTNEIVTLNLYLEPRTYELCVQLVRRGEGEGHPGVPVRRFTTAS 540  
QY 541 IGLPPRGGLNLLPKSQTTLNLTWQPIFPSSSEDDFYVEVERRSVQKSDQONIKVPGNLTSV 600  
DB 541 IGLPPRGGLNLLPKSQTTLNLTWQPIFPSSSEDDFYVEVERRSVQKSDQONIKVPGNLTSV 600  
QY 601 LNNLHPRQYVVRARVNTKAQEKSEDLTANTLSLILPPQENIKISNTHSSAVISWT 660  
DB 601 LNNLHPRQYVVRARVNTKAQEKSEDLTANTLSLILPPQENIKISNTHSSAVISWT 660  
QY 661 ILDGYSISITIRYKVGKNEQHVVDVKIKNATIIQYQLKGLEPETAYQVDIFAENNIGS 720  
DB 661 ILDGYSISITIRYKVGKNEQHVVDVKIKNATIIQYQLKGLEPETAYQVDIFAENNIGS 720  
QY 721 SNPAFSEHVLTPESQAPADLGGKMLLJAILGSAGMTCLTVLLAPLIIILQKRVQR 780  
DB 721 SNPAFSEHVLTPESQAPADLGGKMLLJAILGSAGMTCLTVLLAPLIIILQKRVQR 780  
QY 781 MAQAFONVREBPVQFNSGTLALNRKVNPNPTIYVLDWMDIKFQDVIIGENFGQVLK 840  
DB 781 MAQAFONVREBPVQFNSGTLALNRKVNPNPTIYVLDWMDIKFQDVIIGENFGQVLK 840  
QY 841 ARIKKDGLRMDAAIKRMKEYASKDHRDPAGEVLCKLGHHPNIIINLLGACEHRYLYL 900  
DB 841 ARIKKDGLRMDAAIKRMKEYASKDHRDPAGEVLCKLGHHPNIIINLLGACEHRYLYL 900  
QY 901 ALEYAPHGNLDFLAKSRVLETDPAFANSTASTLSQQLHFAADVARGMDYLSQKQF 960  
DB 901 ALEYAPHGNLDFLAKSRVLETDPAFANSTASTLSQQLHFAADVARGMDYLSQKQF 960  
QY 961 IHRDLAARNILVGENYVAKIADFGLSRQGVVYKKTGRLPVRWMAIESINYSVYTTNSD 1020  
DB 961 IHRDLAARNILVGENYVAKIADFGLSRQGVVYKKTGRLPVRWMAIESINYSVYTTNSD 1020  
QY 1021 VMSYGVLLWEIVSLGTPYCGMTCAELYEKLPQGYRLEKPLNCDDEVDYDLMRQCWEKPY 1080  
DB 1021 VMSYGVLLWEIVSLGTPYCGMTCAELYEKLPQGYRLEKPLNCDDEVDYDLMRQCWEKPY 1080  
QY 1081 ERPSFAQILVSLNRMLEERKTVYNTTLVEKFTYAGIDCSABEAA 1124  
DB 1081 ERPSFAQILVSLNRMLEERKTVYNTTLVEKFTYAGIDCSABEAA 1124  
RESULT 7  
ABU03528  
ID ABU03528 standard; protein; 1124 AA.  
XX  
AC ABU03528;  
XX 21-JAN-2003 (first entry)  
XX  
DE Angiogenesis-associated human protein sequence #73.  
XX  
KW Human; angiogenesis-associated transcript; angiogenesis;  
KW angiogenesis-associated disease; cancer; cytostatic.  
XX  
OS Homo sapiens.

XX WO200279492-A2.  
XX PD 10-OCT-2002.  
XX  
XX 14-FEB-2002; 2002WO-US004915.  
XX  
XX 14-FEB-2001; 2001US-00784356.  
XX PR 22-FEB-2001; 2001US-00791390.  
XX PR 19-APR-2001; 2001US-02854759.  
XX PR 03-AUG-2001; 2001US-03100259.  
XX PR 13-NOV-2001; 2001US-03506669.  
XX PR 29-NOV-2001; 2001US-03342449.  
XX  
XX (EOSB-) EOS BIOTECHNOLOGY INC.  
XX  
XX Murray R, Glynne R, Watson SR, Aziz N;  
XX WPI; 2003-040681/03.  
XX DR N-PSDB; ABX08812.  
XX  
XX Detecting angiogenesis-associated transcript in a cell for diagnosing and  
XX treating cancer by contacting a sample with a polynucleotide that  
XX exhibits changes in expression level as a function of time in tissue  
XX undergoing angiogenesis.  
XX  
XX Example 2; Page 250; 291pp; English.  
XX  
XX The present invention relates to methods and compositions for detecting  
XX an angiogenesis-associated transcript in a cell in a patient. The method  
XX involves contacting a biological sample from the patient with a  
XX polynucleotide that selectively hybridizes to a sequence at least 80%  
XX identical to any of the angiogenesis-associated human polynucleotide  
XX sequences given in the specification. These angiogenesis-associated  
XX polynucleotide sequences comprise genes that exhibit changes in  
XX expression levels as a function of time in tissue undergoing  
XX angiogenesis. The method and the polynucleotide sequences of the  
XX invention are useful for diagnosing and treating angiogenesis and  
XX angiogenesis-associated diseases e.g. cancer. The polynucleotide  
XX sequences are also useful in the gene therapy of such disorders. The  
XX angiogenesis-associated proteins encoded by the polynucleotide sequences  
XX are useful as a vaccine for therapeutic and prophylactic immunisation.  
XX ABU03456-ABU03569 represent angiogenesis-associated protein sequences  
XX  
XX Sequence 1124 AA;  
XX  
XX Query Match 100.0%; Score 6013; DB 6; Length 1124;  
XX Best Local Similarity 100.0%; Pred. No. 0;  
XX Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 1 MDLSASLVLCGVSLLSGTVEGAMDILILNSPLVSDAETSLTCASGWRPHEPITIGRD 60  
XX  
XX 1 MDLSASLVLCGVSLLSGTVEGAMDILILNSPLVSDAETSLTCASGWRPHEPITIGRD 60  
XX  
XX 61 FEALMNQHDPLEVTDVTREWAKVWVKREKASKINGAYFCEGVRGVEAIRITWKMQRQ 120  
XX  
XX 61 FEALMNQHDPLEVTDVTREWAKVWVKREKASKINGAYFCEGVRGVEAIRITWKMQRQ 120  
XX  
XX 121 QASFLPATLTMTVDKGDNNVINSFKVKLIKEDDAVIYKNGSFTHSVPRHEVPDILVHLPH 180  
XX  
XX 121 QASFLPATLTMTVDKGDNNVINSFKVKLIKEDDAVIYKNGSFTHSVPRHEVPDILVHLPH 180  
XX  
XX 181 APODAGVVSARYIGNLTFTSAFTRLIVRCAQKWGPCNHLCTACWNNGVCHSDTGEC 240  
XX  
XX 181 APODAGVVSARYIGNLTFTSAFTRLIVRCAQKWGPCNHLCTACWNNGVCHSDTGEC 240  
XX  
XX 241 ICPPGFMGRTECKACELHTFGRCTKERCSGQGGCKSYVFCPLDPYGCSCATGWKGLQCN 300  
XX  
XX 241 ICPPGFMGRTECKACELHTFGRCTKERCSGQGGCKSYVFCPLDPYGCSCATGWKGLQCN 300  
XX  
XX 301 ACHPGFYGPDCKLRCSNNGENCDRFGQCLCSFGWQGLQCEREGIPRMTPKIVLDPDHE 360  
XX  
XX 301 ACHPGFYGPDCKLRCSNNGENCDRFGQCLCSFGWQGLQCEREGIPRMTPKIVLDPDHE 360

QY 361 VNSGKFNPICKASGWLPTNEMTLVKPDGTVLHPKDFNHTDHFVAIFTIIRILPPDSG 420  
DB 361 VNSGKFNPICKASGWLPTNEMTLVKPDGTVLHPKDFNHTDHFVAIFTIIRILPPDSG 420  
QY 421 VVVCSTVNTVAGWVEKPFNISVKVLPKPLNAPNVIDTGFNFAVINISSBPYFGDGPISKKK 480  
DB 421 VVVCSTVNTVAGWVEKPFNISVKVLPKPLNAPNVIDTGFNFAVINISSBPYFGDGPISKKK 480  
QY 481 LLYKPVNHYEAWQHIOVTNEIVTLNLEPRTEYELCVQLVRRGEGEGHGPVRRFTTAS 540  
DB 481 LLYKPVNHYEAWQHIOVTNEIVTLNLEPRTEYELCVQLVRRGEGEGHGPVRRFTTAS 540  
QY 541 IGLPPRGINLLPKSQTTLNLTWQPIFPSSDEDDFVEVERRSVQSDQONIKVPGNLTSV 600  
DB 541 IGLPPRGINLLPKSQTTLNLTWQPIFPSSDEDDFVEVERRSVQSDQONIKVPGNLTSV 600  
QY 601 LNLNLHPRQYVVRARVNTKAQGEWSEDLTAWTSLDILPPQPENIKISNITHSSAVISWT 660  
DB 601 LNLNLHPRQYVVRARVNTKAQGEWSEDLTAWTSLDILPPQPENIKISNITHSSAVISWT 660  
QY 661 ILDGYSISITIRYKVGKNEHQVDVKIKNATIIYQYQYKGLPEPETAQVDIFAENNTGS 720  
DB 661 ILDGYSISITIRYKVGKNEHQVDVKIKNATIIYQYQYKGLPEPETAQVDIFAENNTGS 720  
QY 721 SNPAFSELVTLPESOAPADILGGGKMLLIATILGSAGMTCLTVLLAFLIILQIKRANVORR 780  
DB 721 SNPAFSELVTLPESOAPADILGGGKMLLIATILGSAGMTCLTVLLAFLIILQIKRANVORR 780  
QY 781 MAQAFQNVREPAVQFNSGTILALNRKVNNDPTTIYPVLDWMDIKFQDVIGEGNFGQVLK 840  
DB 781 MAQAFQNVREPAVQFNSGTILALNRKVNNDPTTIYPVLDWMDIKFQDVIGEGNFGQVLK 840  
QY 841 ARIKDGRLMDAAIKRMKEYASKDHRDPAGELEVCKLGHHPNINILGACEHRYLYL 900  
DB 841 ARIKDGRLMDAAIKRMKEYASKDHRDPAGELEVCKLGHHPNINILGACEHRYLYL 900  
QY 901 AIEYAPHGNLLDFLRKSRVLETPAFATANSTASTLSSQQLLHFAADVARGMDYLSQKF 960  
DB 901 AIEYAPHGNLLDFLRKSRVLETPAFATANSTASTLSSQQLLHFAADVARGMDYLSQKF 960  
QY 961 IHRDLAARNILVGENYVAKIADFLSRGQEVVYKKTGMRLPVRWMAIESLNYSVYTTNSD 1020  
DB 961 IHRDLAARNILVGENYVAKIADFLSRGQEVVYKKTGMRLPVRWMAIESLNYSVYTTNSD 1020  
QY 1021 VMSYGVLLWEIVSLGGTYPYCGMTCALYKLPQGYRLEKPLNCDDEVYDLMRQCWREKPY 1080  
DB 1021 VMSYGVLLWEIVSLGGTYPYCGMTCALYKLPQGYRLEKPLNCDDEVYDLMRQCWREKPY 1080  
QY 1081 ERPSAQILVSLNRMLEERKTVVNTLYEKFTYAGIDCSABEAA 1124  
DB 1081 ERPSAQILVSLNRMLEERKTVVNTLYEKFTYAGIDCSABEAA 1124  
RESULT 8  
ABU07838  
ID ABU07838 standard; protein; 1124 AA.  
XX  
AC ABU07838;  
XX  
DT 10-MAY-2003 (first entry)  
XX  
DE Human Tie receptor tyrosine kinase 2 (Tie 2).  
XX  
KW Cytostatic; vasodilator; antiinflammatory; cardiant; gene therapy;  
KW ligand-receptor binding modulator; ephrin ligand; angiogenesis;  
KW lymphangiogenesis; aberrant Ephrin-Tie biology; cell growth disorder;  
KW cell migration disorder; cell proliferation disorder; neovascularisation;  
KW ischaemia; infarction; tissue graft; transplant; human; tie 2;  
KW tie receptor tyrosine kinase 2; enzyme.  
XX  
OS Homo sapiens.  
XX

PN WO2003004529-A2.  
XX 16-JAN-2003.  
XX 02-JUL-2002; 2002WO-IB002524.  
XX 02-JUL-2001; 2001US-0302960P.  
XX (LICN ) LICENTIA LTD.  
XX Alitalo K, Kubo H;  
XX WPI; 2003-210341/20.  
XX N-PSDB; ABX12539.  
XX  
XX Identifying modulators of binding between a Tie receptor tyrosine kinase  
XX and an Ephrin ligand, useful for promoting neovascularization, comprises  
XX contacting a Tie receptor with an Ephrin in the presence of a putative  
XX modulator.  
XX  
XX Example 1; Page 72-76; 1999p; English.  
XX  
XX The invention describes a method of identifying a modulator of binding  
XX between a Tie receptor tyrosine kinase and an Ephrin ligand. The method  
XX comprises contacting a Tie receptor composition with an Ephrin  
XX composition in the presence and in the absence of a putative modulator  
XX compound, and detecting the binding between Tie receptor and the Ephrin  
XX in the presence and in the absence of the putative modulator. The method  
XX is useful for identifying a modulator of binding between a Tie receptor  
XX tyrosine kinase and an Ephrin ligand. Modulators identified from the  
XX method are useful in modulating angiogenic processes, including  
XX lymphangiogenesis, for treating diseases associated with aberrant Ephrin-  
XX Tie biology, aberrant growth, migration or proliferation of cells that  
XX express a Tie receptor, or for promoting growth of vessel or  
XX neovascularisation (e.g. ischaemic tissue, an infarction, a new or  
XX chronic compound, or a tissue graft or transplant). This is the amino  
XX acid sequence of human Tie receptor tyrosine kinase 2 (Tie 2)  
XX  
XX Sequence 1124 AA;  
XX  
XX Query Match 100.0%; Score 6013; DB 6; Length 1124;  
XX Best Local Similarity 100.0%; Pred. No. 0;  
XX Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 1 MDLSASLVLCGVSLLSGTVEGAMDILINSLPLVSDAETSLTCIASGRWPEHPTIGRD 60  
DB 1 MDLSASLVLCGVSLLSGTVEGAMDILINSLPLVSDAETSLTCIASGRWPEHPTIGRD 60  
QY 61 FEALMNQHDPLEVTDVTRWAKKVVWREKASKINGAYFCEGRVGRGEAIRITWQRQ 120  
DB 61 FEALMNQHDPLEVTDVTRWAKKVVWREKASKINGAYFCEGRVGRGEAIRITWQRQ 120  
QY 121 QASFLPATLTMTVDKGDNNVIFSKVLIKEBDVVIKNGSFTHSVPRHEVPDILEVHLPH 180  
DB 121 QASFLPATLTMTVDKGDNNVIFSKVLIKEBDVVIKNGSFTHSVPRHEVPDILEVHLPH 180  
QY 181 AQPDAGVYSARYIGNLFSTAFTRLIVRRCBAQKWGPCNHLCTACMNGVCHSDTGEC 240  
DB 181 AQPDAGVYSARYIGNLFSTAFTRLIVRRCBAQKWGPCNHLCTACMNGVCHSDTGEC 240  
QY 241 ICPGFMGRTCEKACBLHTFGTKERCSCGQCKSYVFLPDPYGCSCATWKGLOQNE 300  
DB 241 ICPGFMGRTCEKACBLHTFGTKERCSCGQCKSYVFLPDPYGCSCATWKGLOQNE 300  
QY 301 ACHPGFYGPDCIKRCSNNGEMCDRFGQGLCSFGWQGLQCEGREGIPRMTPKIVLDPDHE 360  
DB 301 ACHPGFYGPDCIKRCSNNGEMCDRFGQGLCSFGWQGLQCEGREGIPRMTPKIVLDPDHE 360  
QY 361 VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAIFTIIRLLPPDSG 420  
DB 361 VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAIFTIIRLLPPDSG 420  
QY 421 VWCSTVNTVAGVVEKPFNISVKVLPKPLNAPNVIDTGHNFVAINISSEPFYFGDPIKSKK 480

DB 421 VWCSTVNTVAGVVEKPFNISVKVLPKPLNAPNVIDTGHNFVAINISSEPFYFGDPIKSKK 480  
QY 481 LLYKPVNHYEAWQHIQVTNEIVTLNLEPRTEYELCVLVRREGEHGHGPGVRRRTTAS 540  
DB 481 LLYKPVNHYEAWQHIQVTNEIVTLNLEPRTEYELCVLVRREGEHGHGPGVRRRTTAS 540  
QY 541 IGLPPRGLNLLPKSQTTLLNTWQPIFPSSSEDDFYVEVERRSVQKSDQONIKVPGNLTSV 600  
DB 541 IGLPPRGLNLLPKSQTTLLNTWQPIFPSSSEDDFYVEVERRSVQKSDQONIKVPGNLTSV 600  
QY 601 LLNLHPREQYVVRARVNTKAQGEWSEDLTATLSDILPPOENIKISNTHSSAVISWT 660  
DB 601 LLNLHPREQYVVRARVNTKAQGEWSEDLTATLSDILPPOENIKISNTHSSAVISWT 660  
QY 661 ILDGYSISSITIRYKVGKNEQHVVDVKIKNATIIQYQLKGLEPETAYQVDIFAENNIGS 720  
DB 661 ILDGYSISSITIRYKVGKNEQHVVDVKIKNATIIQYQLKGLEPETAYQVDIFAENNIGS 720  
QY 721 SNPAFSELVTLPESOAPADLGGGKMLLIIAIGSAGMTCLTVLLAFLIILQLKRAVQRR 780  
DB 721 SNPAFSELVTLPESOAPADLGGGKMLLIIAIGSAGMTCLTVLLAFLIILQLKRAVQRR 780  
QY 781 MAQAFQNVREEPAVQFNSGTLALNRKVNNDPDTIYVLDWMDIKFQDVIGEGNFGQVLK 840  
DB 781 MAQAFQNVREEPAVQFNSGTLALNRKVNNDPDTIYVLDWMDIKFQDVIGEGNFGQVLK 840  
QY 841 ARIKDDGLRMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNIINLLGACEHRCGYLYL 900  
DB 841 ARIKDDGLRMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNIINLLGACEHRCGYLYL 900  
QY 901 AIEYAPHGNLLDPLKRSRVLETPDPAFANSTASTLSSQQLLHPAADVARGMDYLSQKQF 960  
DB 901 AIEYAPHGNLLDPLKRSRVLETPDPAFANSTASTLSSQQLLHPAADVARGMDYLSQKQF 960  
QY 961 HRDLAARNILVGENYVAKIADFGLSRGOEVVVKTMGRLPVRWMAIESLNTSVYTTNSD 1020  
DB 961 HRDLAARNILVGENYVAKIADFGLSRGOEVVVKTMGRLPVRWMAIESLNTSVYTTNSD 1020  
QY 1021 VMSYGVLLWEIVSLGGTPYCGMTCAELYEKLPQGYRLEKPLNCDDEVYDLMRQCWREKPY 1080  
DB 1021 VMSYGVLLWEIVSLGGTPYCGMTCAELYEKLPQGYRLEKPLNCDDEVYDLMRQCWREKPY 1080  
QY 1081 ERPSFAQLVSLNRMLEERKTYVNTTLYEKFYAGIDCSAEAAA 1124  
DB 1081 ERPSFAQLVSLNRMLEERKTYVNTTLYEKFYAGIDCSAEAAA 1124  
RESULT 9  
ADD10371  
ID ADD10371 standard; protein; 1124 AA.  
XX  
XX AC ADD10371;  
XX DT 01-JAN-2004 (first entry)  
XX  
XX Human secreted/transmembrane PRO polypeptide #41.  
XX human; secreted protein; transmembrane protein; cardiovascular disorder;  
KW endothelial disorder; angiogenic disorder; myocardial infarction;  
KW cardiac hypertrophy; trauma; cancer; age-related macular degeneration;  
KW angiogenesis; endothelial cell apoptosis; smooth muscle cell growth;  
KW endothelial cell tube formation.  
XX  
XX Homo sapiens.  
XX  
XX US2003105011-A1.  
XX  
XX 05-JUN-2003.  
XX  
XX 16-AUG-2002; 2002US-00223084.  
XX  
XX 15-SEP-2000; 2000US-0232887P.

PR 20-JUN-2001; 2001WO-US019692.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 20-FEB-2002; 2002US-00081056.  
XX  
PA (GETH ) GENENTECH INC.  
XX Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A;  
PI Godowski PJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Stephan JF;  
PI Watanabe CK, Williams PM, Wood WI, Ye W;  
XX  
DR WPI; 2003-810831/76.  
DR N-PSDB; ADD10370.  
XX  
PT New isolated nucleic acid encoding a secreted and transmembrane  
PT polypeptide for treating a cardiovascular, endothelial, or angiogenic  
PT disorder in a mammal, such as cancer or age-related macular degeneration.  
XX  
PS Claim 11; SEQ ID NO 82; 493pp; English.  
XX  
CC The invention relates to an isolated nucleic acid encoding a secreted and  
CC transmembrane polypeptide (PRO). The nucleic acid, a polypeptide encoded  
CC by the nucleic acid, or an agonist or antagonist, is used to treat a  
CC cardiovascular, endothelial, or angiogenic disorder in a mammal,  
CC preferably a human. The human may have suffered a myocardial infarction  
CC or has cardiac hypertrophy, trauma, a cancer, or age-related macular  
CC degeneration. The cardiac hypertrophy is characterised by the presence of  
CC an elevated level of PGP-2 alpha. A PRO polypeptide, given in the  
CC specification, or an agonist is used to inhibit or stimulate endothelial  
CC cell growth in a mammal. PRO21 or an agonist is used to induce cardiac  
CC hypertrophy. PRO1376 or PRO1449 is used to stimulate angiogenesis.  
CC PRO4302 or an agonist is used to induce endothelial cell apoptosis. A PRO  
CC polypeptide, given in the specification, or an agonist is used to  
CC stimulate or inhibit smooth muscle cell growth, or to induce endothelial  
CC cell tube formation. The present sequence represents the amino acid  
CC sequence of a PRO polypeptide of the invention.  
XX  
SQ Sequence 1124 AA;  
  
Query Match 100.0%; Score 6013; DB 7; Length 1124;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MDSLASLVLCGVSLLSGTVEGAMDILLINSIPLVSDAETSLTCTASGWRPHEPTIGRD 60  
DB 1 MDSLASLVLCGVSLLSGTVEGAMDILLINSIPLVSDAETSLTCTASGWRPHEPTIGRD 60  
  
QY 61 FEALMNQHDPLEVTQDVTREWAKKVVWREKASKINGAYFCEGRVGRGAIRIRTMKMRQ 120  
DB 61 FEALMNQHDPLEVTQDVTREWAKKVVWREKASKINGAYFCEGRVGRGAIRIRTMKMRQ 120  
  
QY 121 QASFLPATLTMTVDKDNVNI8FKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
DB 121 QASFLPATLTMTVDKDNVNI8FKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
  
QY 181 AQPQAGVVSARYIIGNLFTSAFTRLIVRCEAQKWPCECNHLCTACMNGVCHEDTGEC 240  
DB 181 AQPQAGVVSARYIIGNLFTSAFTRLIVRCEAQKWPCECNHLCTACMNGVCHEDTGEC 240  
  
QY 241 ICPFGFMGTCTKACELHTFGTCTKERC8G8GCKSVYFCLPDPYGCSCATWGKGLQCN 300  
DB 241 ICPFGFMGTCTKACELHTFGTCTKERC8G8GCKSVYFCLPDPYGCSCATWGKGLQCN 300  
  
QY 301 ACHPGFGPDKCLKSCNNGECDF8QGLCSPGWGLQCEGIPRMTPKIVLDPDHE 360  
DB 301 ACHPGFGPDKCLKSCNNGECDF8QGLCSPGWGLQCEGIPRMTPKIVLDPDHE 360  
  
QY 361 VNSGFNPTICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHF8VAIFTIIRILPPDSG 420  
DB 361 VNSGFNPTICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHF8VAIFTIIRILPPDSG 420  
  
QY 421 VVWCSVNTVAGWVEKPFNISVKVLKPPLNAPNVIDTGHNFVAVINISSEPYFGDGPIKSKK 480  
DB 421 VVWCSVNTVAGWVEKPFNISVKVLKPPLNAPNVIDTGHNFVAVINISSEPYFGDGPIKSKK 480

QY 481 LLYKPVNHYEAWQHIOVTNEIIVTLNLYLPRTTEYELCVOLVREGEGEHGPVRRFTTAS 540  
DB 481 LLYKPVNHYEAWQHIOVTNEIIVTLNLYLPRTTEYELCVOLVREGEGEHGPVRRFTTAS 540  
  
QY 541 IGLPPPRGLNLLPKSOTTINLTWQPIFPSSSEDDFYVEVERRSVQSDQONIKVPGNLTSV 600  
DB 541 IGLPPPRGLNLLPKSOTTINLTWQPIFPSSSEDDFYVEVERRSVQSDQONIKVPGNLTSV 600  
  
QY 601 LLNHLHPREQYVVRARVNTKAGWESEDLTAWTLDSDILPPQPENIKISNITHSSAVISWT 660  
DB 601 LLNHLHPREQYVVRARVNTKAGWESEDLTAWTLDSDILPPQPENIKISNITHSSAVISWT 660  
  
QY 661 ILDGYSISITIRYKVOGKNEQDQVVDVKIKNATIIQYQKGLPETAYQVDI1FAENNIGS 720  
DB 661 ILDGYSISITIRYKVOGKNEQDQVVDVKIKNATIIQYQKGLPETAYQVDI1FAENNIGS 720  
  
QY 721 SNPAFSHELVTLPESQAPADLGGGRMLLIATILGSAGMTCLTVLLAFLIILQLKRANVQR 780  
DB 721 SNPAFSHELVTLPESQAPADLGGGRMLLIATILGSAGMTCLTVLLAFLIILQLKRANVQR 780  
  
QY 781 MAQAFQNVREPAVQFNSGTTALNRKVNNDPTTYPVLDWNDIKFQDVIGEGNFGVLK 840  
DB 781 MAQAFQNVREPAVQFNSGTTALNRKVNNDPTTYPVLDWNDIKFQDVIGEGNFGVLK 840  
  
QY 841 ARIKXGGLRMDAAIKRMKEYASKDHRDPAGLEVLCKLGHHPNIINLLGACEHGRGILYL 900  
DB 841 ARIKXGGLRMDAAIKRMKEYASKDHRDPAGLEVLCKLGHHPNIINLLGACEHGRGILYL 900  
  
QY 901 AIEYAPHGNLLDPLKRSRVLETDPAFAIANSTASTLSSQQLLHPAADVARGMDYLSQKF 960  
DB 901 AIEYAPHGNLLDPLKRSRVLETDPAFAIANSTASTLSSQQLLHPAADVARGMDYLSQKF 960  
  
QY 961 IHRDLAARNILVGENYVAKIADFGLSRGQEVVYKKTGMRLPVRWMAIESLNSVVTNDS 1020  
DB 961 IHRDLAARNILVGENYVAKIADFGLSRGQEVVYKKTGMRLPVRWMAIESLNSVVTNDS 1020  
  
QY 1021 VWSYGVLLWEIVSLGGTYPYCGMTCALYEKLPQGVRLKPLNCDDEVYDLMRQCWREKPY 1080  
DB 1021 VWSYGVLLWEIVSLGGTYPYCGMTCALYEKLPQGVRLKPLNCDDEVYDLMRQCWREKPY 1080  
  
QY 1081 ERPSFAQILVSLNRLMRLEERKTYVNTTLYEKFTYAGIDCSAEAA 1124  
DB 1081 ERPSFAQILVSLNRLMRLEERKTYVNTTLYEKFTYAGIDCSAEAA 1124  
  
RESULT 10  
ADD11331  
ID ADD11331 standard; protein; 1124 AA.  
XX AC ADD11331;  
XX DT 01-JAN-2004 (first entry)  
XX DE Human secreted/transmembrane PRO polypeptide #41.  
XX KW human; secreted protein; transmembrane protein; cardiovascular disorder;  
KW endothelial disorder; angiogenic disorder; myocardial infarction;  
KW cardiac hypertrophy; trauma; cancer; age-related macular degeneration;  
KW angiogenesis; endothelial cell apoptosis; smooth muscle cell growth;  
KW endothelial cell tube formation.  
XX OS Homo sapiens.  
XX PN US2003105013-A1.  
XX PD 05-JUN-2003.  
XX PF 16-AUG-2002; 2002US-00223090.  
XX PR 20-JUN-2001; 2001WO-US019692.  
XX PR 09-JUL-2001; 2001WO-US021735.  
XX PR 20-FEB-2002; 2002US-00081056.

XX (GETH ) GENENTECH INC.  
PA Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A;  
PI Godowski PJ, Gurney AL, Hillan KJ, Mesters SA, Pan J, Stephan JF;  
PI Watanabe CK, Williams PM, Wood WJ, Ye W;  
XX WPI; 2003-801242/75.  
DR N-PSDB; ADD11330.  
XX  
PT New isolated nucleic acid encoding a secreted and transmembrane  
PT polypeptide, useful for treating a cardiovascular, endothelial, or  
PT angiogenic disorder in a mammal, such as cancer or age-related macular  
PT degeneration.  
XX  
PS Claim 11; SEQ ID NO 82; 493pp; English.  
XX  
CC The invention relates to an isolated nucleic acid encoding a secreted and  
CC transmembrane polypeptide (PRO). The nucleic acid, a polypeptide encoded  
CC by the nucleic acid, or an agonist or antagonist, is used to treat a  
CC cardiovascular, endothelial, or angiogenic disorder in a mammal,  
CC preferably a human. The human may have suffered a myocardial infarction  
CC or has cardiac hypertrophy, trauma, a cancer, or age-related macular  
CC degeneration. The cardiac hypertrophy is characterized by the presence of  
CC an elevated level of PGP-2 alpha. A PRO polypeptide, given in the  
CC specification, or an agonist is used to inhibit or stimulate endothelial  
CC cell growth in a mammal. PRO21 or an agonist is used to induce cardiac  
CC hypertrophy. PRO1376 or PRO1449 is used to stimulate angiogenesis.  
CC PRO4302 or an agonist is used to induce endothelial cell apoptosis. A PRO  
CC polypeptide, given in the specification, or an agonist is used to  
CC stimulate or inhibit smooth muscle cell growth, or to induce endothelial  
CC cell tube formation. The present sequence represents the amino acid  
CC sequence of a PRO polypeptide of the invention.  
XX  
SQ Sequence 1124 AA;  
  
Query Match 100.0%; Score 6013; DB 7; Length 1124;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MDLSASLVCGVSLLSGVTEGAMDILINSLPLVSDAETSLTCTASGRWPHPTIGRD 60  
DB 1 MDLSASLVCGVSLLSGVTEGAMDILINSLPLVSDAETSLTCTASGRWPHPTIGRD 60  
  
QY 61 FEALMNQHDPLEVTDQVTRWAKVVKREKASKINGAYFCEGRVGEAIRRTMKRQ 120  
DB 61 FEALMNQHDPLEVTDQVTRWAKVVKREKASKINGAYFCEGRVGEAIRRTMKRQ 120  
  
QY 121 QASFLPATLTMTVDKGDVNIISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
DB 121 QASFLPATLTMTVDKGDVNIISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
  
QY 181 APODAGVVSARYIGNLFTSAFTRLIVRCEAQKWGPCNHLCTACMNGVCHEDTGC 240  
DB 181 APODAGVVSARYIGNLFTSAFTRLIVRCEAQKWGPCNHLCTACMNGVCHEDTGC 240  
  
QY 241 ICPPGFMGTCBACELHTFGRTCKERCSCGQCKSVYFCLPDPYGCSCATWKGILQCN 300  
DB 241 ICPPGFMGTCBACELHTFGRTCKERCSCGQCKSVYFCLPDPYGCSCATWKGILQCN 300  
  
QY 301 ACHPGYGPDCKLRCSCNNGEMCDRPGQCLCSFGWQGLQCEREGIPRMTPKIVLDPDHE 360  
DB 301 ACHPGYGPDCKLRCSCNNGEMCDRPGQCLCSFGWQGLQCEREGIPRMTPKIVLDPDHE 360  
  
QY 361 VNSGKNPICKASGWPLTNEEMTLVKPDGTVLHPKDFNHTDHFSAIETIIRILPPDSG 420  
DB 361 VNSGKNPICKASGWPLTNEEMTLVKPDGTVLHPKDFNHTDHFSAIETIIRILPPDSG 420  
  
QY 421 VVVCSTVNTVAGWVEKPFNISKVLPKPLNAPNVIDTGHNFVAVINISSEPYFGDGPDKSK 480  
DB 421 VVVCSTVNTVAGWVEKPFNISKVLPKPLNAPNVIDTGHNFVAVINISSEPYFGDGPDKSK 480  
  
QY 481 LLYKPVNHYEAWQHIOVTNEIVTLNLYBPRTEYELCVLRREGEGEGHGPVRRFTTAS 540

DB 481 LLYKPVNHYEAWQHIOVTNEIVTLNLYBPRTEYELCVLRREGEGEGHGPVRRFTTAS 540  
QY 541 IGLPPRGLNLLPKSQTTLNLTWQIFPSSSEDDFYVEVERRSVQKSDQONIKVPGNLTSV 600  
DB 541 IGLPPRGLNLLPKSQTTLNLTWQIFPSSSEDDFYVEVERRSVQKSDQONIKVPGNLTSV 600  
QY 601 LLNHLHPREQYVVRVNTKAGWSEDLTAWTSLDILPPQPENIKISNITHSSAVISWT 660  
DB 601 LLNHLHPREQYVVRVNTKAGWSEDLTAWTSLDILPPQPENIKISNITHSSAVISWT 660  
QY 661 ILDGYSISSITIRYKVGKNEQHDVVKIKNATIIQYQKLEPETAYOVDFPAENNIGS 720  
DB 661 ILDGYSISSITIRYKVGKNEQHDVVKIKNATIIQYQKLEPETAYOVDFPAENNIGS 720  
QY 721 SNPAFSHELVTLPESQAPADLGGGKMLLAIILGSAGMTCLTVLLAFLIILQLKRANVQR 780  
DB 721 SNPAFSHELVTLPESQAPADLGGGKMLLAIILGSAGMTCLTVLLAFLIILQLKRANVQR 780  
QY 781 MAQAFQNVREPAVQFNSGTLALNRKVNPPDPTIYVLDVNDIKFQDVIGEGNFGQVLK 840  
DB 781 MAQAFQNVREPAVQFNSGTLALNRKVNPPDPTIYVLDVNDIKFQDVIGEGNFGQVLK 840  
QY 841 ARIKDDGLRMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNIIILGACEHGRGYLYL 900  
DB 841 ARIKDDGLRMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNIIILGACEHGRGYLYL 900  
QY 901 ALEYAPHGNLLDFLRKSVLETPDPAFANSTASTLSQQLLHFAADVARGMDYLSQKQF 960  
DB 901 ALEYAPHGNLLDFLRKSVLETPDPAFANSTASTLSQQLLHFAADVARGMDYLSQKQF 960  
QY 961 IHRDLAARNILVGENYAKIADFGLSRGQEVVYKKTGMRLPVRWMAIESLNTSVYTTNSD 1020  
DB 961 IHRDLAARNILVGENYAKIADFGLSRGQEVVYKKTGMRLPVRWMAIESLNTSVYTTNSD 1020  
QY 1021 VWSYGVLLWEIVSLGTPYCGMTCAELYEKLPGQYRLEKPLNCDDDEVYDLMEQCWREKPY 1080  
DB 1021 VWSYGVLLWEIVSLGTPYCGMTCAELYEKLPGQYRLEKPLNCDDDEVYDLMEQCWREKPY 1080  
QY 1081 ERPSQAQLVLSNRLMBERKTYVNTTLYEKFYAGIDCSAEAA 1124  
DB 1081 ERPSQAQLVLSNRLMBERKTYVNTTLYEKFYAGIDCSAEAA 1124  
  
RESULT 11  
ADD37124  
ID ADD37124 standard; protein; 1124 AA.  
XX  
AC ADD37124;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Human secreted/transmembrane PRO polypeptide #41.  
XX  
KW human; secreted protein; transmembrane protein; cardiovascular disorder;  
KW endothelial disorder; angiogenic disorder; myocardial infarction;  
KW cardiac hypertrophy; trauma; cancer; age-related macular degeneration;  
KW angiogenesis; endothelial cell apoptosis; smooth muscle cell growth;  
KW endothelial cell tube formation.  
XX  
OS Homo sapiens.  
XX  
PN US2003105012-A1.  
XX  
PD 05-JUN-2003.  
XX  
PF 16-AUG-2002; 2002US-00223088.  
XX  
PR 15-SEP-2000; 2000US-0232887P.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 20-FEB-2002; 2002US-00081056.  
XX



(GETH ) GENENTECH INC.

PA Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A; Stephan JF;  
PI Godowski PU, Gurney AL, Hillan KJ, Marsters SA, Pan J, Ye W;  
PI Watanabe CK, Williams PM, Wood WI, Ye W;  
DR WPI; 2003-829354/77.  
XX N-PSDB; ADD37123.

XX New isolated nucleic acids encoding a secreted and transmembrane  
PT polypeptide for treating a cardiovascular, endothelial, or angiogenic  
PT disorder in a mammal, such as cancer or age-related macular degeneration.  
XX  
PS Claim 11; SEQ ID NO 82; 492pp; English.

XX The invention relates to an isolated nucleic acid encoding a secreted and  
CC transmembrane polypeptide (PRO). The nucleic acid, a polypeptide encoded  
CC by the nucleic acid, or an agonist or antagonist, is used to treat a  
CC cardiovascular, endothelial, or angiogenic disorder in a mammal,  
CC preferably a human. The human may have suffered a myocardial infarction  
CC or has cardiac hypertrophy, trauma, a cancer, or age-related macular  
CC degeneration. The cardiac hypertrophy is characterized by the presence of  
CC an elevated level of PGR-2 alpha. A PRO polypeptide, given in the  
CC specification, or an agonist is used to inhibit or stimulate endothelial  
CC cell growth in a mammal. PRO21 or an agonist is used to induce cardiac  
CC hypertrophy. PRO1376 or PRO1449 is used to stimulate angiogenesis.  
CC PRO4302 or an agonist is used to induce endothelial cell apoptosis. A PRO  
CC polypeptide, given in the specification, or an agonist is used to  
CC stimulate or inhibit smooth muscle cell growth, or to induce endothelial  
CC cell tube formation. The present sequence represents the amino acid  
CC sequence of a PRO polypeptide of the invention.

XX Sequence 1124 AA;

Query Match 100.0%; Score 6013; DB 7; Length 1124;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLSASLVLGVSLLSGTVEGAMDILILINSIPLVSDAETSLTCTASGRWPEPTIGRD 60  
DB 1 MDLSASLVLGVSLLSGTVEGAMDILILINSIPLVSDAETSLTCTASGRWPEPTIGRD 60  
QY 61 FEALMNQHDPLEVTQDVTREWAKVWVKREKASKINGAYFCEGVRGVAIRITMKWRQ 120  
DB 61 FEALMNQHDPLEVTQDVTREWAKVWVKREKASKINGAYFCEGVRGVAIRITMKWRQ 120  
QY 121 QASFLPATLTMTVDKGNVNIISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
DB 121 QASFLPATLTMTVDKGNVNIISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
QY 181 AQPDAGVVSARYIGNLFSTAFTRLIVRCEAQKWGPCNHLCTACMNGVCHEDTGEC 240  
DB 181 AQPDAGVVSARYIGNLFSTAFTRLIVRCEAQKWGPCNHLCTACMNGVCHEDTGEC 240  
QY 241 ICPPGFMGRTCEKACELHTFGTCKERCSSGQCKSYVFCPLDPDYGCSATCWKGLQNE 300  
DB 241 ICPPGFMGRTCEKACELHTFGTCKERCSSGQCKSYVFCPLDPDYGCSATCWKGLQNE 300  
QY 301 ACHPGFYGPDCKLRCSCNNGCMCDRFQGCCLSPGQGLQCEREGIPRMTPTKIVDLPDHE 360  
DB 301 ACHPGFYGPDCKLRCSCNNGCMCDRFQGCCLSPGQGLQCEREGIPRMTPTKIVDLPDHE 360  
QY 361 VNSGKFNPICKASGHPLEPNEEMTLVKPDGTVLHPKDNHHTDHFSAFTIHRILPPDSG 420  
DB 361 VNSGKFNPICKASGHPLEPNEEMTLVKPDGTVLHPKDNHHTDHFSAFTIHRILPPDSG 420  
QY 421 VVWCSVNTVAGNVEKPFNISVSKPLNAPNVIDTGNFVAINISSPEYFGDGPIKSKK 480  
DB 421 VVWCSVNTVAGNVEKPFNISVSKPLNAPNVIDTGNFVAINISSPEYFGDGPIKSKK 480  
QY 481 LLYKPNVHYEAWQHIQVNTNEIVTLNLYLEPRTYEYLCVQLVRRGEGEGHPGVRFTTAS 540  
DB 481 LLYKPNVHYEAWQHIQVNTNEIVTLNLYLEPRTYEYLCVQLVRRGEGEGHPGVRFTTAS 540

QY 541 IGLPPRGINLLPKSQTTLNLTWQPIFPSSSEDDFYVEVERBSVQKSDOONIKVPGNLTSV 600  
DB 541 IGLPPRGINLLPKSQTTLNLTWQPIFPSSSEDDFYVEVERBSVQKSDOONIKVPGNLTSV 600  
QY 601 LIANNLHPREQYVVRVNTKAGWSEDLTAWTSLDILPPQPENIKISNITHSSAVISWT 660  
DB 601 LIANNLHPREQYVVRVNTKAGWSEDLTAWTSLDILPPQPENIKISNITHSSAVISWT 660  
QY 661 ILDGYSISSITIRYKVGKNEHQVDVKIKNATIIYQYQKGLEPETAYQVDIFAENNIGS 720  
DB 661 ILDGYSISSITIRYKVGKNEHQVDVKIKNATIIYQYQKGLEPETAYQVDIFAENNIGS 720  
QY 721 SNPAFSHELVTLPESQAPADILGGGKMLLIALILGSAGMTCLTVLLAFLIILQLKRANVORR 780  
DB 721 SNPAFSHELVTLPESQAPADILGGGKMLLIALILGSAGMTCLTVLLAFLIILQLKRANVORR 780  
QY 781 MAQAFQNVREPAVQFNSGTALANRKVNKNPDPTIYPVLDMNDIKFQDVIGEGNFGQVLK 840  
DB 781 MAQAFQNVREPAVQFNSGTALANRKVNKNPDPTIYPVLDMNDIKFQDVIGEGNFGQVLK 840  
QY 841 ARIKDGRLMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNIINLLGACEHRGYLYL 900  
DB 841 ARIKDGRLMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNIINLLGACEHRGYLYL 900  
QY 901 AIEYAPHGNLLDLFLKRSRVLETDPAFATANSTASTLSSQQLLHPAADVARGMDVLSQKF 960  
DB 901 AIEYAPHGNLLDLFLKRSRVLETDPAFATANSTASTLSSQQLLHPAADVARGMDVLSQKF 960  
QY 961 IHRDLAARNILVGENYVAKIADFGLSRGQEVVVKTMGRLPVRWMAIESLNSYVVTNSD 1020  
DB 961 IHRDLAARNILVGENYVAKIADFGLSRGQEVVVKTMGRLPVRWMAIESLNSYVVTNSD 1020  
QY 1021 VMSYGVLLWEIVSLGGTPTCYGTCALYKELPQGYRLEKPLNCCDDDEVYDLMRQCWRKPY 1080  
DB 1021 VMSYGVLLWEIVSLGGTPTCYGTCALYKELPQGYRLEKPLNCCDDDEVYDLMRQCWRKPY 1080  
QY 1081 ERPSAQILVSLNRMLEBKTYVNTTLYEKFTYAGIDCSABEAA 1124  
DB 1081 ERPSAQILVSLNRMLEBKTYVNTTLYEKFTYAGIDCSABEAA 1124

RESULT 12

ADF45095

ID ADF45095 standard; protein; 1124 AA.

XX ADF45095;

XX 12-FEB-2004 (first entry)

XX Human kinase TIE2.

XX Human; protein kinase; enzyme; inhibitor; TIE2.

XX Homo sapiens.

XX WO2003081210-A2.

XX 02-OCT-2003.

XX 20-MAR-2003; 2003WO-US008725.

XX 21-MAR-2002; 2002US-0366892P.

XX (SUNE-) SUNESIS PHARM INC.

XX Prescott JC, Braisted A;

XX WPI; 2003-865136/80.

XX Identifying ligand binding to inactive conformation of target protein  
PT kinase (T) comprises contacting the conformation modified (T) which  
PT contains reactive group at binding site, with ligands and detecting



PT	kinase-ligand conjugate formation.
XX	Disclosure; SEQ ID NO 64; 260pp; English.
PS	The present invention relates to a method for identifying a ligand (L),
XX	which binds to an inactive conformation of target protein kinase (T). The
CC	method involves contacting inactive conformation of (T), which contains
CC	or is modified to contain a reactive group at or near a binding site of
CC	interest, with one or more ligand candidates capable of covalently
CC	bonding to the reactive group thus forming a kinase-(L) conjugate (C).
CC	The method is useful for identifying protein kinase inhibitors that
CC	preferentially bind to inactive conformation of a target protein kinase.
CC	The present sequence is a protein kinase which may be modified via an
CC	amino acid substitution, for use in the method of the invention.
XX	Sequence 1124 AA;
SQ	Query Match 100.0%; Score 6013; DB 7; Length 1124;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MDSLASLVLCGVSLLSGTVEGAMDILINSLPLVSDAETSLTCTASGWRPHEPITIGRD 60
DB	1 MDSLASLVLCGVSLLSGTVEGAMDILINSLPLVSDAETSLTCTASGWRPHEPITIGRD 60
QY	61 FEALMNQHDPLEVTDVTRWAKVWVKREKASKINGAYFCEGRVVRGEAIRITWKMQRQ 120
DB	61 FEALMNQHDPLEVTDVTRWAKVWVKREKASKINGAYFCEGRVVRGEAIRITWKMQRQ 120
QY	121 QASFLPATLTMTVDGDNVNI SFKKVLIKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180
DB	121 QASFLPATLTMTVDGDNVNI SFKKVLIKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180
QY	181 AQPQAGVYSARYIGNLFTSAFTRLIVRCAQKMGPECNHLCTACMNGVCHEDTGB 240
DB	181 AQPQAGVYSARYIGNLFTSAFTRLIVRCAQKMGPECNHLCTACMNGVCHEDTGB 240
QY	241 ICPPGFMGTRCKACELHFTGTRCKERCQSGCKSYVFCCLPDPYGCSCATWKGQCN 300
DB	241 ICPPGFMGTRCKACELHFTGTRCKERCQSGCKSYVFCCLPDPYGCSCATWKGQCN 300
QY	301 ACHPGFYGPDCIKRSCNNGENCDRFOGCLCSFGWQGLQCEGREGIPRMTPKIVLDPD 360
DB	301 ACHPGFYGPDCIKRSCNNGENCDRFOGCLCSFGWQGLQCEGREGIPRMTPKIVLDPD 360
QY	361 VNSGKFNPTCKASGWLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAI FTIHRILPPDSG 420
DB	361 VNSGKFNPTCKASGWLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAI FTIHRILPPDSG 420
QY	421 VVVCVNTVAGVVEKPFNISVKVLPKPLNAPNVIDTGHNFVAVINISSEPYFGDGP 480
DB	421 VVVCVNTVAGVVEKPFNISVKVLPKPLNAPNVIDTGHNFVAVINISSEPYFGDGP 480
QY	481 LLYKPNHYEAOHQVTHETI VTLNLEPRTYELCVQLVRGEGEHGPVRFTTAS 540
DB	481 LLYKPNHYEAOHQVTHETI VTLNLEPRTYELCVQLVRGEGEHGPVRFTTAS 540
QY	541 IGLPPRGNLNLPKSTOTLNLNLTWQPIFPSSDDFYVEVERRSVQKSDQONIKVPGNLT 600
DB	541 IGLPPRGNLNLPKSTOTLNLNLTWQPIFPSSDDFYVEVERRSVQKSDQONIKVPGNLT 600
QY	601 LLNHLHPREQYVVRVNTKAGWSEDLTAWTSLDILPPQENIKISNITHSSAVISMT 660
DB	601 LLNHLHPREQYVVRVNTKAGWSEDLTAWTSLDILPPQENIKISNITHSSAVISMT 660
QY	661 ILDGYSISSITIRYKVGKNEOHVDVKIKNATIIQYQLKGLBETAYQVDI FAENNIGS 720
DB	661 ILDGYSISSITIRYKVGKNEOHVDVKIKNATIIQYQLKGLBETAYQVDI FAENNIGS 720
QY	721 SNPAFSEHVLTPESQAPADLGGKMLLAIILGSAWMTCLTVLLAFLIILQLKRVQR 780
DB	721 SNPAFSEHVLTPESQAPADLGGKMLLAIILGSAWMTCLTVLLAFLIILQLKRVQR 780

QY	781 MAQAFQNVREBPAPQVNSGTTALNRKVKNPDPPTIYPVLDMNDIKFQDVIGSGNFQVLK 840
DB	781 MAQAFQNVREBPAPQVNSGTTALNRKVKNPDPPTIYPVLDMNDIKFQDVIGSGNFQVLK 840
QY	841 ARIKXGRLMDAAIRKMEYASKDHRDPAGLEVLCKLGHHPNIIINLIGACEHRYLYL 900
DB	841 ARIKXGRLMDAAIRKMEYASKDHRDPAGLEVLCKLGHHPNIIINLIGACEHRYLYL 900
QY	901 AIEYAPHGNLLDFLRKSRVLETDPAPAFIANSTASTLSSQQLLHFAADVARGMDYLSQKQF 960
DB	901 AIEYAPHGNLLDFLRKSRVLETDPAPAFIANSTASTLSSQQLLHFAADVARGMDYLSQKQF 960
QY	961 IHRDLAARNILVGENYVAKIADFGLSRGQEVVYKKTGMRLPVRWMAIESLNTSVYTTNSD 1020
DB	961 IHRDLAARNILVGENYVAKIADFGLSRGQEVVYKKTGMRLPVRWMAIESLNTSVYTTNSD 1020
QY	1021 VMSYGVLLWEIVSLGTPYCGMTCAELYEKLPGQVRLKPLNCDDEVDVDMRCQWREKPY 1080
DB	1021 VMSYGVLLWEIVSLGTPYCGMTCAELYEKLPGQVRLKPLNCDDEVDVDMRCQWREKPY 1080
QY	1081 ERPSAQILVSLNRMLERKTYVNTTLYEKFTYAGIDCSAEBA 1124
DB	1081 ERPSAQILVSLNRMLERKTYVNTTLYEKFTYAGIDCSAEBA 1124
RESULT 13	
ADE41332	
ID	ADE41332 standard; protein; 1124 AA.
XX	
AC	ADE41332;
XX	
DT	29-JAN-2004 (first entry)
XX	
DE	Human secreted/transmembrane PRO polypeptide #41.
XX	
KW	human; secreted protein; transmembrane protein; cardiovascular disorder;
KW	endothelial disorder; angiogenic disorder; myocardial infarction;
KW	cardiac hypertrophy; trauma; cancer; age-related macular degeneration;
KW	angiogenesis; endothelial cell apoptosis; smooth muscle cell growth;
XX	endothelial cell tube formation.
OS	Homo sapiens.
XX	
PN	US2003100497-A1.
XX	
PD	29-MAY-2003.
XX	
PF	16-AUG-2002; 2002US-00223085.
XX	
PR	20-JUN-2001; 2001WO-US019692.
PR	09-JUL-2001; 2001WO-US021735.
PR	20-FEB-2002; 2002US-00081056.
XX	
PA	(GETH ) GENENTECH INC.
XX	
PI	Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A;
PI	Godowski PJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Stephan JF;
PI	Watanabe CK, Williams PM, Wood WI, Ye W;
XX	
DR	WPI; 2004-008957/01.
DR	N-PSDB; ADE41331.
XX	
PT	New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO205 or
PT	PRO214, useful in molecular biology, chromosome and gene mapping, in
PT	generating antisense RNA and DNA, and for treating disorders involving
PT	angiogenesis.
XX	
PS	Claim 11; SEQ ID NO 82; 492pp; English.
XX	
CC	The invention relates to an isolated nucleic acid encoding a secreted and
CC	transmembrane polypeptide (PRO). The nucleic acid, a polypeptide encoded
CC	by the nucleic acid, or an agonist or antagonist, is used to treat a
CC	cardiovascular, endothelial, or angiogenic disorder in a mammal,

CC preferably a human. The human may have suffered a myocardial infarction  
 CC or has cardiac hypertrophy, trauma, a cancer, or age-related macular  
 CC degeneration. The cardiac hypertrophy is characterised by the presence of  
 CC an elevated level of PGR-2 alpha. A PRO polypeptide, given in the  
 CC specification, or an agonist is used to inhibit or stimulate endothelial  
 CC cell growth in a mammal. PRO21 or an agonist is used to induce cardiac  
 CC hypertrophy. PRO1376 or PRO1449 is used to stimulate angiogenesis.  
 CC PRO4302 or an agonist is used to induce endothelial cell apoptosis. A PRO  
 CC polypeptide, given in the specification, or an agonist is used to  
 CC stimulate or inhibit smooth muscle cell growth, or to induce endothelial  
 CC cell tube formation. The present sequence represents the amino acid  
 CC sequence of a PRO polypeptide of the invention.

XX Sequence 1124 AA;

Query Match 100.0%; Score 6013; DB 8; Length 1124;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLSASLVLCGVSLLSGTVEGAMDILINSLPLVSDAETSLTCTASGRPHEPITIGRD 60  
 DB 1 MDLSASLVLCGVSLLSGTVEGAMDILINSLPLVSDAETSLTCTASGRPHEPITIGRD 60  
 QY 61 FEALMNQHDPLEVTQDVTREWAKVWVKREKASKINGAYFCEGRVGEAIRIRTMKMRQ 120  
 DB 61 FEALMNQHDPLEVTQDVTREWAKVWVKREKASKINGAYFCEGRVGEAIRIRTMKMRQ 120  
 QY 121 QASFLPATLTMTVDKGDNNISFKKVLKEBDAYIKNGSFIHSPRHEVPDILEVHLPH 180  
 DB 121 QASFLPATLTMTVDKGDNNISFKKVLKEBDAYIKNGSFIHSPRHEVPDILEVHLPH 180  
 QY 181 APOQAGVVSARYIGNLFTSAFTRLIVRCEAQKGPCECNHLCACMNNGVCHSDTGC 240  
 DB 181 APOQAGVVSARYIGNLFTSAFTRLIVRCEAQKGPCECNHLCACMNNGVCHSDTGC 240  
 QY 241 ICPPGFMGRTCEKACELHTFGRTKERCSSQGCKSYVFCCLPDPYGCSCATWGKLGQNE 300  
 DB 241 ICPPGFMGRTCEKACELHTFGRTKERCSSQGCKSYVFCCLPDPYGCSCATWGKLGQNE 300  
 QY 301 ACHPGFYGDCKLRCSCNNGECNDRFQGLCSFGWGLQCEGREGIPRMTPKIVLDPDHE 360  
 DB 301 ACHPGFYGDCKLRCSCNNGECNDRFQGLCSFGWGLQCEGREGIPRMTPKIVLDPDHE 360  
 QY 361 VNSGKFNPICKASGNPLPTNEBMTLVKPDGTVLHPKDNHTDHFSAVFTIHRILPPDSG 420  
 DB 361 VNSGKFNPICKASGNPLPTNEBMTLVKPDGTVLHPKDNHTDHFSAVFTIHRILPPDSG 420  
 QY 421 VVWCSVNTVAGMVEKPFNISVKVLPKPLNAPNVIDTGHNFVINISSBPYFGDGPISKK 480  
 DB 421 VVWCSVNTVAGMVEKPFNISVKVLPKPLNAPNVIDTGHNFVINISSBPYFGDGPISKK 480  
 QY 481 LLYKPNVHYEAWQHIVQNTNEIIVTLNLEPRTEYELCVQLVRRGEGBGHPGVRPFTAS 540  
 DB 481 LLYKPNVHYEAWQHIVQNTNEIIVTLNLEPRTEYELCVQLVRRGEGBGHPGVRPFTAS 540  
 QY 541 IGLPPRGNLPLPKSQTLLNTWQPIFPSSDDFYVEVERRSVQKSDQONIKVPGNLTSV 600  
 DB 541 IGLPPRGNLPLPKSQTLLNTWQPIFPSSDDFYVEVERRSVQKSDQONIKVPGNLTSV 600  
 QY 601 LNNLHPREQYVVRARVNTKAQWSEDLTAWTSLDILPPQENIKISNITHSSAVISWT 660  
 DB 601 LNNLHPREQYVVRARVNTKAQWSEDLTAWTSLDILPPQENIKISNITHSSAVISWT 660  
 QY 661 ILDGYSISITIRYKQGNEDQHDVVKIKNATIIQYQLKLEPETAYQVDIFAENNTGS 720  
 DB 661 ILDGYSISITIRYKQGNEDQHDVVKIKNATIIQYQLKLEPETAYQVDIFAENNTGS 720  
 QY 721 SNPAFSELVTLPEAOADLGGKMLLIALIGSAGMCTCLTVLLAFLIQLKRVNORR 780  
 DB 721 SNPAFSELVTLPEAOADLGGKMLLIALIGSAGMCTCLTVLLAFLIQLKRVNORR 780  
 QY 781 MAQAFQNVREPAVQFNSGTTALNRKVNPNDDPTIYPVLDMNDIKFQDVGISGNFGQVLK 840  
 DB 781 MAQAFQNVREPAVQFNSGTTALNRKVNPNDDPTIYPVLDMNDIKFQDVGISGNFGQVLK 840

DB 781 MAQAFQNVREPAVQFNSGTTALNRKVNPNDDPTIYPVLDMNDIKFQDVGISGNFGQVLK 840  
 QY 841 ARIKDDGLRMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNIIINLLGACEHRYLYL 900  
 DB 841 ARIKDDGLRMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNIIINLLGACEHRYLYL 900  
 QY 901 AIEYAPHGNLIDFLRKSRVLETDPAPAJANSTASTLSSQQLLHFAADVARGMDYLSQKQF 960  
 DB 901 AIEYAPHGNLIDFLRKSRVLETDPAPAJANSTASTLSSQQLLHFAADVARGMDYLSQKQF 960  
 QY 961 IHRDLAARNILVGENYAKIADFGLSRGOEVVYKVTMGRLPVRWMAIESLNYSVYTTNSD 1020  
 DB 961 IHRDLAARNILVGENYAKIADFGLSRGOEVVYKVTMGRLPVRWMAIESLNYSVYTTNSD 1020  
 QY 1021 VMSYGVILWEIVSLGGTTCGTCALYEKLPQGYRLEKPLNCDDDEVYDLMRQCVRKPY 1080  
 DB 1021 VMSYGVILWEIVSLGGTTCGTCALYEKLPQGYRLEKPLNCDDDEVYDLMRQCVRKPY 1080  
 QY 1081 ERPSAQIILVSLNRMLEBRKTVVNTTLYEKFTYAGIDCSAEAA 1124  
 DB 1081 ERPSAQIILVSLNRMLEBRKTVVNTTLYEKFTYAGIDCSAEAA 1124

RESULT 14  
 ADH43515  
 ID ADH43515 standard; protein; 1124 AA.  
 XX AC ADH43515;  
 XX DT 25-MAR-2004 (first entry)  
 XX DE Human PRO polypeptide #41.  
 XX KW Human; PRO; cardiovascular disorder; endothelial disorder;  
 KW angiogenic disorder; endothelial cell growth; cardiac hypertrophy;  
 KW cell apoptosis; cell tube formation; angiogenesis;  
 KW smooth muscle cell growth; myocardial infarction; trauma; cancer;  
 KW age-related macular degeneration; cytostatic; cardiant;  
 KW cerebroprotective; ophthalmological; vulneryary.  
 XX OS Homo sapiens.  
 XX PN US2003224984-A1.  
 XX PD 04-DEC-2003.  
 XX PF 26-NOV-2002; 2002US-00305654.  
 XX PR 20-JUN-2001; 2001WO-US019692.  
 XX PA (GETH ) GENENTECH INC.  
 XX PI Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A;  
 PI Godowski PJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Paoni NF;  
 PI Stephan JF, Watanabe CK, Williams PM, Wood WI, Ye W;  
 XX WPI: 2004-042166/04.  
 DR N-PSDB; ADH43514.  
 XX  
 PT New PRO polypeptides and nucleic acids encoding the polypeptides, useful  
 PT for treating myocardial infarction, cardiac hypertrophy, trauma, cancer,  
 PT or age-related macular degeneration.  
 XX  
 PS Claim 11; SEQ ID NO 82; 492pp; English.  
 XX  
 CC The invention relates to human PRO polypeptides and the PRO  
 CC polynucleotides encoding them. The invention also relates to treating  
 CC cardiovascular, endothelial or angiogenic disorders in mammals,  
 CC inhibiting endothelial cell growth, stimulating endothelial cell growth,  
 CC inducing cardiac hypertrophy, cell apoptosis or cell tube formation and  
 CC stimulating angiogenesis or smooth muscle cell growth by administering  
 CC polypeptides of the invention. The PRO polypeptides and polynucleotides  
 CC are useful for treating cardiovascular, endothelial or angiogenic

CC	disorders, e.g. myocardial infarction, cardiac hypertrophy, trauma, cancer or age-related macular degeneration. The PRO polynucleotides are useful as hybridization probes in chromosome and gene mapping and in generating antisense RNA and DNA, and for chromosome identification and tissue typing. The PRO polypeptides and polynucleotides are also useful in gene therapy and as molecular weight markers for protein electrophoresis purposes. This sequence represents a human PRO polypeptide of the invention.	
XX	Sequence 1124 AA;	
SQ	Query Match 100.0%; Score 6013; DB 8; Length 1124; Best Local Similarity 100.0%; Pred. No. 0; Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MDLSASLVLCGVSLLSGTVEGAMDLILINSPLVSDAETSLTCIASGRWPHETIGRD 60	
DB	1 MDLSASLVLCGVSLLSGTVEGAMDLILINSPLVSDAETSLTCIASGRWPHETIGRD 60	
QY	61 FEALMNHQODPLEVTDVTRWAKVVMKREKASKINGAYFCGEVRVGEAIRITWKMQR 120	
DB	61 FEALMNHQODPLEVTDVTRWAKVVMKREKASKINGAYFCGEVRVGEAIRITWKMQR 120	
QY	121 QASPLPATLTMTVDKGNVNI8FKVLIKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180	
DB	121 QASPLPATLTMTVDKGNVNI8FKVLIKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180	
QY	181 APOQDAGVVSARYIGGNLFTSAFTRLIVRCEAQKWPCECNHLCACMNGVCHEDTGC 240	
DB	181 APOQDAGVVSARYIGGNLFTSAFTRLIVRCEAQKWPCECNHLCACMNGVCHEDTGC 240	
QY	241 ICPPGFMGRCTCBKACELHFTGRTCKERCQGBGCKSYVFCCLPDPYGCSCATGKGLQCN 300	
DB	241 ICPPGFMGRCTCBKACELHFTGRTCKERCQGBGCKSYVFCCLPDPYGCSCATGKGLQCN 300	
QY	301 ACHPGFYGPDCIKRCSNNGENCDRFOGCLCSPGWQGLQCEBEGIPRMTPKIVLDPDHE 360	
DB	301 ACHPGFYGPDCIKRCSNNGENCDRFOGCLCSPGWQGLQCEBEGIPRMTPKIVLDPDHE 360	
QY	361 VNSGKFNPICKASGWPLPNEEMTLVKPDGTVLHPKDFNHTDHF5VAIFTIHRILPPDSG 420	
DB	361 VNSGKFNPICKASGWPLPNEEMTLVKPDGTVLHPKDFNHTDHF5VAIFTIHRILPPDSG 420	
QY	421 VVVCSTVNTVAGVVEKPFNISVKVLPKPLNAPNVIDTGHNFVAVINISSEPYFGDGPISKK 480	
DB	421 VVVCSTVNTVAGVVEKPFNISVKVLPKPLNAPNVIDTGHNFVAVINISSEPYFGDGPISKK 480	
QY	481 LLYKPNHYEAOHQIQTNEIIVTLNLEPRTEYELCVQVRGEGEGHPGVRPRTTAS 540	
DB	481 LLYKPNHYEAOHQIQTNEIIVTLNLEPRTEYELCVQVRGEGEGHPGVRPRTTAS 540	
QY	541 IGLPPRGLNLLPKSOTTLNLTWQPIFPPSSEDDFYVEVERRSVQKSDQONIKVPGNLTSV 600	
DB	541 IGLPPRGLNLLPKSOTTLNLTWQPIFPPSSEDDFYVEVERRSVQKSDQONIKVPGNLTSV 600	
QY	601 LNNLHPRQYVVRARVNTKACQSESDTATWLTSDILPPQENIKISNITHSSAVISMT 660	
DB	601 LNNLHPRQYVVRARVNTKACQSESDTATWLTSDILPPQENIKISNITHSSAVISMT 660	
QY	661 ILDGYSISITIRYKVGKNEOHVDVKIKNATIIQYQLKGLEPETAQYVDI1FAENNIGS 720	
DB	661 ILDGYSISITIRYKVGKNEOHVDVKIKNATIIQYQLKGLEPETAQYVDI1FAENNIGS 720	
QY	721 SNPAFSEHVLTPESQAPADLGGKMLLIALIGSAGMTCLTVLLAFLIILQLKRVQR 780	
DB	721 SNPAFSEHVLTPESQAPADLGGKMLLIALIGSAGMTCLTVLLAFLIILQLKRVQR 780	
QY	781 MAQAFQNVREPAVQNSGTLALNRKVNPNPTIYVLDWMDIKFQDVIGEGNFQVLK 840	
DB	781 MAQAFQNVREPAVQNSGTLALNRKVNPNPTIYVLDWMDIKFQDVIGEGNFQVLK 840	
QY	841 ARIKDDGLRMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNIIINLLGACEHGRGILYL 900	

DB	841 ARIKDDGLRMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNIIINLLGACEHGRGILYL 900	
QY	901 ALEYAPHGNLLDLFLRKSRLVLETDPAFAIANSTASTLSQQLLHFAADVARGMDYLSQKQF 960	
DB	901 ALEYAPHGNLLDLFLRKSRLVLETDPAFAIANSTASTLSQQLLHFAADVARGMDYLSQKQF 960	
QY	961 IHRDLAARNILVGENYVAKIADFGLSRGQEVVYKKTGMRLPVRWMAIESLNTSVYTTNSD 1020	
DB	961 IHRDLAARNILVGENYVAKIADFGLSRGQEVVYKKTGMRLPVRWMAIESLNTSVYTTNSD 1020	
QY	1021 VMSYGVLLWEIVSLGTPYCGMTCAELYEKLPGYRLEKPLNCDDEVDYDMRQCWREKY 1080	
DB	1021 VMSYGVLLWEIVSLGTPYCGMTCAELYEKLPGYRLEKPLNCDDEVDYDMRQCWREKY 1080	
QY	1081 ERPSFAQILVSLNRMLEERKTYVNTLYEKFTYAGIDCSAEAA 1124	
DB	1081 ERPSFAQILVSLNRMLEERKTYVNTLYEKFTYAGIDCSAEAA 1124	
RESULT 15		
ADK82860		
ID	ADK82860 standard; protein; 1124 AA.	
XX		
AC	ADK82860;	
XX		
DT	20-MAY-2004 (first entry)	
XX		
DE	Human PRO polypeptide #41.	
XX		
KW	Human; PRO; cardiovascular disorder; endothelial disorder; angiogenic disorder; endothelial cell growth; cardiac hypertrophy; cell apoptosis; cell tube formation; angiogenesis;	
KW	smooth muscle cell growth; myocardial infarction; trauma; cancer;	
KW	age-related macular degeneration; cytostatic; cardiant;	
KW	cerebroprotective; ophthalmological; vulnery.	
XX		
OS	Homo sapiens.	
XX		
PN	US2004043927-A1.	
PD		
XX		
XX	04-MAR-2004.	
PF		
XX	20-FEB-2002; 2002US-00081056.	
PR	19-SEP-1997; 97US-00933821.	
PR	29-JAN-1998; 98US-00015089.	
PR	10-SEP-1998; 98WO-US018824.	
PR	14-SEP-1998; 98WO-US019177.	
PR	16-SEP-1998; 98WO-US019330.	
PR	17-SEP-1998; 98WO-US019437.	
PR	19-NOV-1998; 98US-00180997.	
PR	01-DEC-1998; 98WO-US025108.	
PR	22-DEC-1998; 98US-00218517.	
PR	05-JAN-1999; 99WO-US000106.	
PR	12-FEB-1999; 99US-00214186.	
PR	03-MAR-1999; 99US-00254311.	
PR	08-MAR-1999; 99WO-US005028.	
PR	09-MAR-1999; 99US-00254460.	
PR	12-MAR-1999; 99US-00267213.	
PR	12-APR-1999; 99US-00284291.	
PR	02-JUN-1999; 99WO-US012252.	
PR	14-JUN-1999; 99US-00332928.	
PR	25-AUG-1999; 99US-00380138.	
PR	25-AUG-1999; 99US-00380139.	
PR	01-SEP-1999; 99WO-US020111.	
PR	08-SEP-1999; 99WO-US020594.	
PR	13-SEP-1999; 99WO-US020944.	
PR	15-SEP-1999; 99WO-US021090.	
PR	15-SEP-1999; 99WO-US021547.	
PR	05-OCT-1999; 99WO-US023089.	
PR	15-OCT-1999; 99US-00403154.	
PR	18-OCT-1999; 99US-00403296.	

```
PR 18-OCT-1999; 99US-00403297.
PR 10-NOV-1999; 99US-00423741.
PR 12-NOV-1999; 99US-00423843.
PR 12-NOV-1999; 99US-00423844.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030999.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 23-AUG-2000; 2000WO-US023328.
PR 24-AUG-2000; 2000WO-US023328.
PR 18-SEP-2000; 2000US-00664610.
PR 18-SEP-2000; 2000US-00665350.
PR 08-NOV-2000; 2000US-00709238.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 05-APR-2001; 2001US-00828366.
PR 25-MAY-2001; 2001US-00865028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021086.
PR 09-JUL-2001; 2001WO-US021735.
PR 30-JUL-2001; 2001US-00918585.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 13-AUG-2001; 2001US-00929404.
PR 16-AUG-2001; 2001US-00931836.
PR 28-AUG-2001; 2001US-00941992.
PR 04-SEP-2001; 2001US-00946374.
PR 07-SEP-2001; 2001US-00948901.
PR 15-NOV-2001; 2001US-0002796.
PR 30-NOV-2001; 2001US-0001054.
PR 06-DEC-2001; 2001US-00068867.
PR 19-DEC-2001; 2001US-00028072.
PR 15-JAN-2002; 2002US-00052586.
PR 17-JAN-2002; 2002US-00053107.

PR 18-JAN-2002; 2002US-00052594.
PR 08-FEB-2002; 2002US-00072068.
PR 09-APR-2002; 2002US-00119480.
XX (GETH ) GENENTECH INC.
XX Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A,
PI Godowski PJ, Gurney AL, Hillan KJ, Marsters SA, Pan J, Stephan JF;
PI Watanabe CK, Williams PM, Wood WI, Ye W;
XX WPI; 2004-225727/21.
DR N-PSDB; ADK82859.
XX
XX New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO20080 or
PT PRO2183j, useful in molecular biology, chromosome and gene mapping, in
PT generating antisense RNA and DNA, and in gene therapy.
XX
PS Claim 11; SEQ ID NO 82; 494pp; English.
XX
CC The invention relates to human PRO polypeptides and the PRO
CC polynucleotides encoding them. The invention also relates to treating
CC cardiovascular, endothelial or angiogenic disorders in mammals,
CC inhibiting endothelial cell growth, stimulating endothelial cell growth,
CC inducing cardiac hypertrophy, cell apoptosis or cell tube formation and
CC stimulating angiogenesis or smooth muscle cell growth by administering
CC polypeptides of the invention. The PRO polypeptides and polynucleotides
CC are useful for treating cardiovascular, endothelial or angiogenic
CC disorders, e.g. myocardial infarction, cardiac hypertrophy, trauma,
CC cancer or age-related macular degeneration. The PRO polynucleotides are
CC useful as hybridisation probes in chromosome and gene mapping and in
CC generating antisense RNA and DNA, and for chromosome identification and
CC tissue typing. The PRO polypeptides and polynucleotides are also useful
CC in gene therapy and as molecular weight markers for protein
CC electrophoresis purposes. This sequence represents a human PRO
CC polypeptide of the invention. Note: The sequence data for this patent is
CC also available in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 1124 AA;
Query Match 100.0%; Score 6013; DB 8; Length 1124;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDSLASLVLCGVSLLSGTVGAMDLILINSILPLVSDAETSLTCTASGWRPHEPTIGRD 60
DB 1 MDSLASLVLCGVSLLSGTVGAMDLILINSILPLVSDAETSLTCTASGWRPHEPTIGRD 60
QY 61 FEALMNQHDPLEVTQDVTREWAKKVVWKRKASKINGAYFCEGRVGEAIRIRTMKMRQ 120
DB 61 FEALMNQHDPLEVTQDVTREWAKKVVWKRKASKINGAYFCEGRVGEAIRIRTMKMRQ 120
QY 121 QASFLPATLTWTDKGDNNVTSFKKVLIKEDAVIYKNGSFTHSVPRHEVPDILVHLPH 180
DB 121 QASFLPATLTWTDKGDNNVTSFKKVLIKEDAVIYKNGSFTHSVPRHEVPDILVHLPH 180
QY 181 AQPQDAGVYSARYIGGNLFTSAFTLRVRCBAQKWGPECNHLCCTACWNGVCHDGTGEC 240
DB 181 AQPQDAGVYSARYIGGNLFTSAFTLRVRCBAQKWGPECNHLCCTACWNGVCHDGTGEC 240
QY 241 ICPGFMGRTCEKACELHTFGRTCKERCSCGEGCKSVYVFCPLDPYGCSCATGKGLQCN 300
DB 241 ICPGFMGRTCEKACELHTFGRTCKERCSCGEGCKSVYVFCPLDPYGCSCATGKGLQCN 300
QY 301 ACHPGFYGPDCKLRCSNNGEMCDRFGQCLSPGWQGLQCREREGIPRMTPKIVLDPDHIE 360
DB 301 ACHPGFYGPDCKLRCSNNGEMCDRFGQCLSPGWQGLQCREREGIPRMTPKIVLDPDHIE 360
QY 361 VNSGKFNPICKASGWPPLTNEEMTLVKPDGTVLHPKDFNHTDHFVAFITTHRIILPPDSG 420
DB 361 VNSGKFNPICKASGWPPLTNEEMTLVKPDGTVLHPKDFNHTDHFVAFITTHRIILPPDSG 420
QY 421 VWCVSNTVAGVKEKPFNISVKVLPKPLNAPNVIDTGHNFVINITSBPYFGDGPISKKK 480
```

[illegible]

Search completed: March 14, 2006, 01:43:41  
Job time : 200 secs

This Page Blank (uspto)

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2006, 01:43:59 ; Search time 49 Seconds  
(without alignments)  
2207.095 Million cell updates/sec

Title: US-10-633-742-8  
Perfect score: 6013  
Sequence: 1 MDSLASLVLCGVSLLSGTV.....TTLVEKFTYAGIDCSAREAA 1124

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 80:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6013	100.0	1124	1 I58388	protein-tyrosine k
2	5732.5	95.3	1125	1 S57846	protein-tyrosine k
3	5598.5	93.1	1123	1 JN0712	protein-tyrosine k
4	5580	92.8	1122	2 I54237	protein-tyrosine k
5	5371.5	89.3	1125	1 JH0771	protein-tyrosine k
6	2526	42.0	1136	1 S57845	protein-tyrosine k
7	2515	41.8	1134	1 JN0711	protein-tyrosine k
8	2511	41.8	1138	1 S24066	protein-tyrosine k
9	719.5	12.0	806	2 A35963	protein-tyrosine k
10	717.5	11.9	974	1 A49714	protein-tyrosine k
11	713.5	11.9	797	2 S3579	fibroblast growth
12	703.5	11.7	822	2 S29840	fibroblast growth
13	701.5	11.7	832	2 JH0393	fibroblast growth
14	698.5	11.6	800	2 A40991	heparin-binding gr
15	698.5	11.6	822	1 TVNSFG	fibroblast growth
16	696.5	11.6	800	1 TVHU2P	fibroblast growth
17	696.5	11.6	822	2 I49289	fibroblast growth
18	694.5	11.5	822	1 TVHUFG	fibroblast growth
19	693	11.5	806	1 TVHUFG	fibroblast growth
20	692	11.5	729	2 A49120	fibroblast growth
21	690	11.5	818	2 JC4058	fibroblast growth
22	686	11.4	812	1 A34477	fibroblast growth
23	683	11.4	822	2 S19947	fibroblast growth
24	683	11.4	822	2 B49151	fibroblast growth
25	682.5	11.4	819	1 TVCHFG	fibroblast growth
26	681	11.3	801	2 I55363	fibroblast growth
27	678	11.3	729	2 A56795	fibroblast growth
28	676	11.2	799	2 S18209	fibroblast growth
29	673	11.2	814	1 A39752	fibroblast growth

30	671	11.2	733	2 I49293	fibroblast growth
31	671	11.2	1064	1 S57450	protein-tyrosine k
32	670	11.1	829	2 JC4583	fibroblast growth
33	667.5	11.1	816	2 A49151	fibroblast growth
34	666.5	11.1	650	1 JC1450	fibroblast growth
35	666	11.1	707	2 A38429	keratinocyte growth
36	665.5	11.1	822	2 B54846	fibroblast growth
37	662.5	11.0	823	2 B35963	protein-tyrosine k
38	661	11.0	802	1 TVHUFG	fibroblast growth
39	661	11.0	824	2 S36439	fibroblast growth
40	660.5	11.0	707	2 A54846	fibroblast growth
41	660.5	11.0	1040	2 A57638	receptor tyrosine
42	659	11.0	748	2 S41050	fibroblast growth
43	659	11.0	769	2 S16236	fibroblast growth
44	659	11.0	822	2 A45081	fibroblast growth
45	658.5	11.0	820	2 S17295	fibroblast growth

ALIGNMENTS

RESULT 1

I58388  
protein-tyrosine kinase (EC 2.7.1.112), receptor type tek precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: I58388  
R:Ziegler, S.F.; Bird, T.A.; Schneringer, J.A.; Schooley, K.A.; Baum, P.R.  
Oncogene 8, 663-670, 1993  
A:Title: Molecular cloning and characterization of a novel receptor protein tyrosine kinase  
A:Reference number: I58388; MUID:93173509; PMID:8382358  
A:Accession: I58388  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1124 <RES>  
A:Cross-references: UNIPROT:Q02763; UNIPARC:UPI0000032E58; GB:L06139; NID:9292823; PIDN:  
C:Genetics:  
A:Gene: GDB:TEK  
A:Cross-references: GDB:344185; OMIM:600221  
A:Map position: 9p21-9p21  
C:Function:  
A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
C:Superfamily: protein-tyrosine kinase, receptor type tie; EGF homology; fibronectin type  
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphot  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-1124/Product: protein-tyrosine kinase, receptor type tek #status predicted <MAT>  
F:37-104/Domain: immunoglobulin homology <IM1>  
F:135-137/Region: cell attachment (R-G-D) motif  
F:211-251/Domain: EGF homology <EG1>  
F:255-298/Domain: EGF homology <EG2>  
F:302-340/Domain: EGF homology <EG3>  
F:364-426/Domain: immunoglobulin homology <IM2>  
F:447-527/Domain: fibronectin type III repeat homology <FN3A>  
F:542-625/Domain: fibronectin type III repeat homology <FN3B>  
F:638-720/Domain: fibronectin type III repeat homology <FN3C>  
F:752-772/Domain: transmembrane #status predicted <TM>  
F:822-1099/Domain: protein kinase homology <KIN>  
F:830-838/Region: protein kinase ATP-binding motif  
F:140,158,399,438,464,560,596,649,691/Binding site: carbohydrate (Asn) (covalent) #status  
F:855,872,964/Active site: Lys, Glu, Asp #status predicted





Db 421 VVWCSVNTVSGMVEKPFNISVKVLPKLNAPKVIDTGHNFANINISSEPFYFGDPKSKK 480  
QY 481 LLYKPNVHYEAWQHIOVTNEIVTLNLEPRTEYELCVQLVRRGEGEGHPGVPVRRFTTAS 540  
Db 481 LLYKPNVHYEAWRHIOVTNEIVTLNLEPRTEYELCVQLVRRGEGEGHPGVPVRRFTTAS 540  
QY 541 IGLPPRGVNLPLPKSQTTLNLWQPIFPSSSEDDFYVEVRRSVQ-KSDQONIKVPGNLTLS 599  
Db 541 IGLPPRGVNLPLPKSQTTLNLWQPIFPSSSEDDFYVEVRRSVQMSDQONIKVPGNLTLS 600  
QY 600 VLLNNLHPREQVVRARVNTKAGWSEDLTAWLSDILPPROPENIKISNIHTSSAVISM 659  
Db 601 VLLNNLHPREQVVRARVNTKAGWSEDLTAWLSDILPPROPENIKISNIHTSSAVISM 660  
QY 660 TILDGYSISITIRYKVOGKNEQDQVVDVKIKNATIIYQQLKGLPEPTAYQVDFIANNIG 719  
Db 661 TILDGYSISAIIRYKVOGKNEQDQVVDVKIKNATIIYQQLKGLPEPTAYQVDFIANNIG 720  
QY 720 SNNPAPSHLVLPESQAPADLGGGKMLLIALIGSAGMTCLTVLLAFLIILQKRAVQR 779  
Db 721 SNNPTSSHELTTLSQAPADLGGGKMLLIALIGSAGMTCLTVLLAFLIILQKRAVQR 780  
QY 780 RMAQAFQNVREBPVQFNSGTALNRKAKNPNPTIYPVLDWMDIKPQDVIGEGNFGQVL 839  
Db 781 RMAQAFQNVREBPVQFNSGTALNRKAKNPNPTIYPVLDWMDIKPQDVIGEGNFGQVL 840  
QY 840 KARIKKGRLMDAAIKRMEYASKODHRDPAGELEVLCKLGHHPNIIINLLGACEHRYLY 899  
Db 841 KARIKKGRLMDAAIKRMEYASKODHRDPAGELEVLCKLGHHPNIIINLLGACEHRYLY 900  
QY 900 LAIEYAPHGNLLDFLRKSVLETPDPAFAJANSTASTLSSQQLLHFAADVARGMDYLSQK 959  
Db 901 LAIEYAPHGNLLDFLRKSVLETPDPAFAJANSTASTLSSQQLLHFAADVARGMDYLSQK 960  
QY 960 FTHRDLAARNILVGENYAKIADGRLSGQEVYVKTWGLPVRWMAIESLANSYVYTINS 1019  
Db 961 FTHRDLAARNILVGENYAKIADGRLSGQEVYVKTWGLPVRWMAIESLANSYVYTINS 1020  
QY 1020 DWYSGVLLWEIVSLGTPYCGMTCAELYEKLPGQYRLEKPLNCDDVYDLMRQCWREKP 1079  
Db 1021 DWYSGVLLWEIVSLGTPYCGMTCAELYEKLPGQYRLEKPLNCDDVYDLMRQCWREKP 1080  
QY 1080 YERPSPAQILVSLNRLMERKTYVNTLYEKFYAGIDCSAEAA 1124  
Db 1081 YERPSPAQILVSLNRLMERKTYVNTLYEKFYAGIDCSAEAA 1125

RESULT 3  
JN0712  
protein-tyrosine kinase (BC 2.7.1.112), receptor type tek precursor - mouse  
N:Alternate names: protein-tyrosine kinase, receptor type tie-2  
C:Species: Mus musculus (house mouse)  
C:Date: 27-Jun-1994 #sequence revision 27-Jun-1994 #text\_change 09-Jul-2004  
C:Accession: JN0712; S57848; S43495; S43494; S33142  
R:Irwana, A.; Hanaguchi, I.; Hashiyama, M.; Murayama, Y.; Yasunaga, K.; Suda, T.  
Biochem. Biophys. Res. Commun. 195, 301-309, 1993  
A:Title: Molecular cloning and characterization of mouse TIE and TEK receptor tyrosine kinase  
A:Reference number: JN0711; MUID:93371421; PMID:8395828  
A:Accession: JN0712  
A:Molecule type: mRNA  
A:Residues: 1-1123 <IWA>  
A:Cross-references: UNIPROT:Q02858; UNIPARC:UPI00001572PB  
R:Sato, T.N.; Qin, Y.; Kozak, C.A.; Audus, K.L.  
Proc. Natl. Acad. Sci. U.S.A. 90, 9355-9358, 1993  
A:Title: tie-1 and tie-2 define another class of putative receptor tyrosine kinase genes  
A:Reference number: S57845; MUID:94022374; PMID:8415706  
A:Accession: S57848  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-786,788-1123 <SAT>  
A:Cross-references: UNIPARC:UPI000002998B; EMBL:X71426; NID:g296612; PIDN:CAAS0557.1; PI  
R:Dumont, D.J.

submitted to the EMBL Data Library, April 1993  
A:Reference number: S43495  
A:Accession: S43495  
A:Molecule type: mRNA  
A:Residues: 1-537, 'C', 539-735, 'G', 737-786, 788-1123 <DUM>  
A:Cross-references: UNIPARC:UPI0000146DB3; EMBL:X67553; NID:g297158; PIDN:CAA47857.1; PI  
R:Dumont, D.J.; Yanaguchi, T.P.; Conlon, R.A.; Rossant, J.; Breitman, M.L.  
Oncogene 7, 1471-1480, 1992  
A:Title: tek, a novel tyrosine kinase gene located on mouse chromosome 4, is expressed in  
A:Reference number: S43494; MUID:92334855; PMID:1630810  
A:Accession: S43494  
A:Molecule type: mRNA  
A:Residues: 823-1123 <DUW>  
A:Cross-references: UNIPARC:UPI0000143147; EMBL:X67553  
C:Comment: Receptor tyrosine kinase-ligand systems play an important role in the constit  
C:Genetics:  
A:Map position: 4  
C:Superfamily: protein-tyrosine kinase, receptor type tie; EGF homology; fibronectin type  
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphot  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-1123/Product: protein-tyrosine kinase, receptor type tie-2 #status predicted <NAT>  
F:37-104/Domain: immunoglobulin homology <IM1>  
F:135-137/Region: cell attachment (R-G-D) motif  
F:211-251/Domain: EGF homology <EG1>  
F:255-298/Domain: EGF homology <EG2>  
F:302-340/Domain: EGF homology <EG3>  
F:364-426/Domain: immunoglobulin homology <IM2>  
F:447-525/Domain: fibronectin type III repeat homology <FN3A>  
F:540-624/Domain: fibronectin type III repeat homology <FN3B>  
F:637-719/Domain: fibronectin type III repeat homology <FN3C>  
F:751-771/Domain: transmembrane #status predicted <TMW>  
F:821-1098/Domain: protein kinase homology <KIN>  
F:829-837/Region: protein kinase ATP-binding site  
F:140, 158, 399, 438, 464, 558, 648, 690/Binding site: carbohydrate (Asn) (covalent) #status  
F:854, 871, 963/Active site: Lys, Glu, Asp #status predicted

Query Match 93.1%; Score 5598.5; DB 1; Length 1123;  
Best Local Similarity 92.8%; Pred. No. 1.1e-215;  
Matches 1044; Conservative 37; Mismatches 41; Indels 3; Gaps 2;

QY 1 MDLSASLVLCVSLLSGTTVEGAMDLLILNSIPLVSDAETSLTCTASGWRHEPITIGRD 60  
Db 1 MDLSAGLVLCVSLLSGTTVEGAMDLLILNSIPLVSDAETSLTCTASGWRHEPITIGRD 60  
QY 61 FEALMNQHPDPLEVTQDVTREWAKKVVWREKASKINGAYFCGEGRVGRGAIIRTKMRQ 120  
Db 61 FEALMNQHPDPLEVTQDVTREWAKKVVWREKASKINGAYFCGEGRVGRGAIIRTKMRQ 120  
QY 121 QASFLPATLTMTVDKGDVNIISFKKVLKEEDAVIYKNGSFTHSVPRHVPDILEVHLPH 180  
Db 121 QASFLPATLTMTVDKGDVNIISFKKVLKEEDAVIYKNGSFTHSVPRHVPDILEVHLPH 180  
QY 181 AQPQAGVYSARYIGGNLFTSAFTLIVRRCBAQKWPCECNHLCTACMNGVCHEDTGEC 240  
Db 181 AQPQAGVYSARYIGGNLFTSAFTLIVRRCBAQKWPCECNHLCTACMNGVCHEDTGEC 240  
QY 241 ICPPGFMGRCTCKACELHTFGRTCKERCSCGQGCCKSYVFCCLPDPYGCSCATGWKGLQCN 300  
Db 241 ICPPGFMGRCTCKACELHTFGRTCKERCSCGQGCCKSYVFCCLPDPYGCSCATGWKGLQCN 300  
QY 301 ACHPGFYGDCKLRCSNNGEMCDRFQGLCSFGWQGLQCEGREGIPRMTPKIVLDLPHIE 360  
Db 301 ACHPGFYGDCKLRCSNNGEMCDRFQGLCSFGWQGLQCEGREGIPRMTPKIVLDLPHIE 360  
QY 361 VNSGKFNPTCKASGWPPLNNEEMTLVKDGTVLHPKDFNHTDHFVAIFTHRIILPPDSG 420  
Db 361 VNSGKFNPTCKASGWPPLNNEEMTLVKDGTVLHPKDFNHTDHFVAIFTHRIILPPDSG 420  
QY 421 VVWCSVNTVAGVVEKPFNISVKVLPKLNAPNVIDTGHNFANINISSEPFYFGDPKSKK 480  
Db 421 VVWCSVNTVAGVVEKPFNISVKVLPKLNAPNVIDTGHNFANINISSEPFYFGDPKSKK 480  
QY 481 LLYKPNVHYEAWQHIOVTNEIVTLNLEPRTEYELCVQLVRRGEGEGHPGVPVRRFTTAS 540

Db	481	LFYKPVN--QAWKYLEVTNEIFTLNLYLPRDTEYELCVQLARPEGEHGPVRRFTTAS	538
Qy	541	IGLPPRGLNLLPKSQTLLNTWQPIFPSSSEDDFFVEVERRSVQ-KSDQONIKVPGNLTS	599
Db	539	IGLPPRGLSLLPKSQTALNTWQPIFTNSDEDFYVEVERRSLOQTSQONIKVPGNLTS	598
Qy	600	VLLNHLPREQVVRVAVNTKAQGSWSDLTAWTLSDILPPPENIKISNITHSSAVISW	659
Db	599	VLLSNLVPREQVTVARVNTKAQGSWSELRAWTLSDILPPPENIKISNITDSTAMVSW	658
Qy	660	TILDGYSISSITIRYKVGQKNEHDQVVKIKNATIIYQLKLEPETAYQVDIFAENNIG	719
Db	659	TIVDGYSSISSIIIRYKVGQKNEHDQVVKIKNATVYQLKLEPETYYHVDIFAENNIG	718
Qy	720	SSNPAFSLHVTLPESQADPAGGKMLLIIALIGSAGMTCTVLLAFILIIQLKRAVQR	779
Db	719	SSNPAFSLHRTLPHSPASADILGGKMLLIIALIGSAGMTCTVLLAFILMLQLKRAVQR	778
Qy	780	RMAQAFQNVREEPVAFQNSGTLALNRKKNPDPTIYPVLDWNDIKFDQVIGEGNFGQVL	839
Db	779	RMAQAFQNVREEPVAFQNSGTLALNRKKNPDPTIYPVLDWNDIKFDQVIGEGNFGQVL	838
Qy	840	KARIKDGRLMDAAIKRMKEYASKDDHRDFAGELEVLCKLGHHPNIINLLGACEHRGVL	899
Db	839	KARIKDGRLMDAAIKRMKEYASKDDHRDFAGELEVLCKLGHHPNIINLLGACEHRGVL	898
Qy	900	LAIEYAPHGNLLDFURKSRVLETPAPAIANSTASTLSQQLLHFAADVARGMDYLSQKQ	959
Db	899	LAIEYAPHGNLLDFURKSRVLETPAPAIANSTASTLSQQLLHFAADVARGMDYLSQKQ	958
Qy	960	FTHRDLAARNILVGENYVAKTADFGLSRGQEVYVKTGMRLPVRWMAIESLNSYVYTTNS	1019
Db	959	FTHRDLAARNILVGENYVAKTADFGLSRGQEVYVKTGMRLPVRWMAIESLNSYVYTTNS	1018
Qy	1020	DWMSGVLLWEIVSLGGTPYCGMTCAELYEKLPGQYRLEKPLNCDDVYDLMRCQWREKP	1079
Db	1019	DWMSGVLLWEIVSLGGTPYCGMTCAELYEKLPGQYRLEKPLNCDDVYDLMRCQWREKP	1078
Qy	1080	VERPSFAQILVSLNRLBERKTYNTTYEKFYAGIDCSAEEA	1124
Db	1079	VERPSFAQILVSLNRLBERKTYNTTYEKFYAGIDCSAEEA	1123

RESULT 4

154237

protein-tyrosine kinase (SC 2.7.1.112), receptor type tek precursor - mouse

N:Alternate names: protein-tyrosine kinase receptor type tie2

C:Species: Mus sp. (mouse)

C>Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 21-Jan-2000

C:Accession: I54237

R:Kuntzing, A.S.; Stacker, S.A.; Wilks, A.F.

Growth Factors 9, 99-105, 1993

A:Title: tie2, a putative protein tyrosine kinase from a new class of cell surface recep

A:Reference number: I54237; MUID:94031116; PMID:8217221

A:Accession: I54237

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1122 <RES>

A:Cross-references: UNIPARC:UPI0000170C78; GB:S67051; NID:g452873; PIDN:AAB28663.1; PID:

C:Genetics:

A:Gene: tie2

C:Superfamily: protein-tyrosine kinase, receptor type tie; EGF homology; fibronectin typ

C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho

F:37-104/Domain: immunoglobulin homology <IMM>

F:211-251/Domain: EGF homology <EG1>

F:255-298/Domain: EGF homology <EG2>

F:302-340/Domain: EGF homology <EG3>

F:540-624/Domain: fibronectin type III repeat homology <3FR>

F:820-1097/Domain: protein kinase homology <KIN>

F:828-836/Region: protein kinase ATP-binding motif

Query Match 92.8%; Score 5580; DB 2; Length 1122;

		Best Local Similarity 92.6%; Pred. No. 6e-215;	
		Matches 1042; Conservative 37; Mismatches 42; Indels 4; Gaps 3;	
Qy	1	MDSLASLVLCVGSLLSGTVGAMDLIIILNSLPLVSDAETSLTCTIASGRWHPHEPITIGRD	60
Db	1	MDSLAGLVLCVGSLLLYGVWGAMDLIIILNSLPLVSDAETSLTCTIASGWHPEPITIGRD	60
Qy	61	FEALMNQHDPLEVTQDVTREWAKVWKRKASKINGAYCEGVRGQATIRTMKWRQ	120
Db	61	FEALMNQHDPLEVTQDVTREWAKVWKRKASKINGAYCEGVRGQATIRTMKWRQ	120
Qy	121	QASFLPATLTWTVDKGNVNISFKKVLIIKEBDVVIYKNGSFIHSPRHEVPDILEVHLPH	180
Db	121	QASFLPATLTWTVDKGNVNISFKKVLIIKEBDVVIYKNGSFIHSPRHEVPDILEVHLPH	180
Qy	181	AQPDAGVYSARYIGGNLFTSAFTRLIVRRCEAQKGPBECNHLCTACMNNGVCHBDTGEC	240
Db	181	AQPDAGVYSARYIGGNLFTSAFTRLIVRRCEAQKGPDCSRPCTTCKNNGVCHBDTGEC	240
Qy	241	ICPPGFMGRTCEKACELHTFGRTCKERCSGOGCKSVYVCLPDPYGCSCATGWGLQONE	300
Db	241	ICPPGFMGRTCEKACEPHTFGRTCKERCSGPEGCKSVYVCLPDPYGCSCATGWGLQONE	300
Qy	301	ACHPGFYGPDCKLRCSNNGEMCDRFQGLCSPGWQGLQCEREGIPRMTPKIVDLPDHIE	360
Db	301	ACPSGYGPDCKLRCHCTNBEICDRFQGLCSQGWGLQCEKEGRPMTPQILEDLPDHIE	360
Qy	361	VNSGKFNPICKASGWPLPTNBEMLVKPDGTVLHPKDPNHTDHFVAIFTTHRIILPPDSG	420
Db	361	VNSGKFNPICKASGWPLPTNBEMLVKPDGTVLHPKDPNHTDHFVAIFTVNRVLPPDSG	420
Qy	421	VWVCSVNTVAGWKEKPFNISVKVLPKPLNAPNVDTHNEFVAINISSEPSYFGDGIKSKK	480
Db	421	VWVCSVNTVAGWKEKPFNISVKVLPEPLHAPNVVDTHNFALINISSEPSYFGDGIKSKK	480
Qy	481	LLYKPVNHYEAWQHIQVTNEIVTLNLYLPRTEYELCVQLVRRGEGEHGPVRRFTTAS	540
Db	481	LFYKPVN--QAWKYLEVTNEIFTLNLYLPRDTEYELCVQLARPEGEHGPVRRFTTAS	538
Qy	541	IGLPPRGLNLLPKSQTLLNTWQPIFPSSSEDDFFVEVERRSVQ-KSDQONIKVPGNLTS	599
Db	539	IGLPPRGLSLLPKSQTALNTWQPIFTNSDEDFYVEVERRSLOQTSQONIKVPGNLTS	598
Qy	600	VLLNHLPREQVVRVAVNTKAQGSWSDLTAWTLSDILPPPENIKISNITHSSAVISW	659
Db	599	VLLSNLVPREQVTVARVNTKAQGSWSELRAWTLSDILPPPENIKISNITDSTAMVSW	658
Qy	660	TILDGYSISSITIRYKVGQKNEHDQVVKIKNATIIYQLKLEPETAYQVDIFAENNIG	719
Db	659	TIVDGYSSISSIIIRYKVGQKNEHDQVVKIKNATVYQLKLEPETYYHVDIFAENNIG	718
Qy	720	SSNPAFSLHVTLPESQADPAGGKMLLIIALIGSAGMTCTVLLAFILIIQLKRAVQR	779
Db	719	SSNPAFSLHRTLPHSPASADILGGKMLLIIALIGSAGMTCTVLLAFILMLQLKRAVQR	778
Qy	780	RMAQAFQNVREEPVAFQNSGTLALNRKKNPDPTIYPVLDWNDIKFDQVIGEGNFGQVL	839
Db	779	RMAQAFQNVREEPVAFQNSGTLALNRKKNPDPTIYPVLDWNDIKFDQVIGEGNFGQVL	837
Qy	840	KARIKDGRLMDAAIKRMKEYASKDDHRDFAGELEVLCKLGHHPNIINLLGACEHRGVL	899
Db	838	KARIKDGRLMDAAIKRMKEYASKDDHRDFAGELEVLCKLGHHPNIINLLGACEHRGVL	897
Qy	900	LAIEYAPHGNLLDFURKSRVLETPAPAIANSTASTLSQQLLHFAADVARGMDYLSQKQ	959
Db	898	LAIEYAPHGNLLDFURKSRVLETPAPAIANSTASTLSQQLLHFAADVARGMDYLSQKQ	957
Qy	960	FTHRDLAARNILVGENYVAKTADFGLSRGQEVYVKTGMRLPVRWMAIESLNSYVYTTNS	1019
Db	958	FTHRDLAARNILVGENYVAKTADFGLSRGQEVYVKTGMRLPVRWMAIESLNSYVYTTNS	1017
Qy	1020	DWMSGVLLWEIVSLGGTPYCGMTCAELYEKLPGQYRLEKPLNCDDVYDLMRCQWREKP	1079









Db 126 SPSSGDDDDDDDES-----EDTGVPFWTRDPKMEKLLAVPAANTVRFCRCPAGNPTPTIY 181  
QY 506 YLEPRTYELCVQLVRRGGEGHCPVRRFTTASIGLPPRGLNLLPKSQTTMLTWOP 565  
Db 182 WLKNGKEPK-----GHRIG-----GKIL-----RHQWESLWMS 211  
QY 566 IFPSDEDDFYVEVER-----SVQSDQONI-----KVPGNLTSLVNLN----- 605  
Db 212 VVPSDRGNYTCVENKYGNIHRTYQLDVLERSPHRPILQAGLPANQTVVVGNSVNEFFCHKV 271  
QY 606 ----HPREQYVVRARVNTKAOCEWSEDLTAWTLDLPPQENIKISNITHSSAVISWTI 661  
Db 272 YSDAQPHQWLKHVEVNGSKYG-----PDGTPYV-----TV 302  
QY 662 LDGYSISSITIRYKQGN-EDQHDV-KIKNATIIQYOLKLEPETAYQVDIFANNIG 719  
Db 303 L-----KTAGVNTTQKEILEILYLRNVT-----EDAGEYTCLAGNSIG 340  
QY 720 SNNPAFSH-----LVTLPEQAADLGGKMLLIALILGSGAGTCLTVLLAFLLIQLKRN 776  
Db 341 ----FSHSAWLTVPABELMEMDDSGSVYAGILSYGTGLVFLVILVIVIVICRMKPN 395  
QY 777 VQRMAQAFQNV-----REPAVQFNSG-----TLALNRKVKGNP 811  
Db 396 KXAMNTTVQKVKPKLKKQVLSLSSNMNNTPLVTRILSSSDGPMANVSLELPP 455  
QY 812 DPTIYPLVDW-----NDIKPDQVIGBNFGQVLKAR-----IKKD-----GLRMDAAIKRMKEYAS 862  
Db 456 DP-----KWELARSRLTGKPLGEGCGQVVAEAGIDKDKPKNKAITVAVKMLKDDAT 509  
QY 863 KDDHRDFAGELEVLCGLGHHPNINLLGACEHRGYLYIAIEVAPHGNLLDFURKSRVLET 922  
Db 510 KDXLSDLVSEMEMMKWIGKHNIINLLGACTQDGLYVLVEYASKNREYLRRARPPCM 569  
QY 923 DPAPAIANSTASTLSQQLHFAADVARGMDVLSQKFTHRDLAARNLILVGENYAKIAD 982  
Db 570 DYSFDTCKLPERQLTFKDLVSCAYQVARGMEYLASQKCIHRDLAARNVLVETEDNVMKIAD 629  
QY 983 FGLSR-----GOEYVVKTKMRLPVRMAIESLNVSVTTNSDVSYSYGLLWEIVSLGCTPY 1039  
Db 630 FGLARDVENIDYKKTNGRLPVKMAPEALFDRVYTHQSDVMSFGVLLWEIFTLGGSPY 689  
QY 1040 CQMTCAELYEKLPQGYRLEKPLNCDDVYDLRMQCRKPKYRPPSPAQILVSLNRLSER 1099  
Db 690 PGIPVEELFKLLKEGRMDKPNCTHLYIMRECWHAVPSQRTFKQLVEDLDRVLWT 749  
QY 1100 KT--YVNTTL-YEKFTYAGID-----CSA 1120  
Db 750 STDEYLDLSVPFEQYSPAGQDTHSTCSS 777

RESULT 10  
A49714  
protein-tyrosine kinase (EC 2.7.1.112) c-eyk precursor - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 05-Oct-2004  
C:Accession: A49714  
R:Jia, R.; Hanafusa, H.  
J. Biol. Chem. 269, 1839-1844, 1994  
A:Title: The proto-oncogene of v-eyk (v-ryk) is a novel receptor-type protein tyrosine k  
A:Reference number: A49714; MUID:94124527; PMID:7507487  
A:Accession: A49714  
A:Status: preliminary  
A:Molecule type: mRNA  
A:References: 1-974 <JIA>  
A:Cross-references: UNIPROT:Q90777; UNIPARC:UPI00000PCCB3; GB:I21719; NID:G438522; PIDN:  
C:Superfamily: Tyrosine-protein kinase, receptor proto-oncogene AXL type; fibronectin ty  
C:Keywords: ATP; glycoprotein; phosphotransferase; transmembrane protein; tyrosine-speci  
F:201-254/Domain: immunoglobulin homology <IM>  
F:575-851/Domain: protein kinase homology <KIN>  
F:583-591/Domain: protein kinase ATP-binding motif

Query Match 11.9%; Score 717.5; DB 1; Length 974;

Best Local Similarity 28.0%; Pred. No. 1.2e-21;  
Matches 249; Conservative 141; Mismatches 328; Indels 172; Gaps 41;  
QY 330 LCSQWQGLQEREGIP--RMTPKVLDLPDRIEVNSGKFNPI-----CKASGWPLPT 379  
Db 44 LARPLWAKHHRPKGLTSGRWPPQAG-PSATSVQQLKFNPTVGHVINELKDVFNCSI 102  
QY 380 NEEMTLVKP-----DGTVLHPKDFNHTDHFV-----AIFTIHRILPPDSGV 421  
Db 103 KVPQLLVFPDPSGILWKDGRHLHTLDRIATSHFBLDEEVAMTSTFSIRAAQSDNGS 162  
QY 422 WVCSVNTVAG--MVEKPFENISVKVLPKPLNAPNVIDTGHNFAVINISSEPYFGDGPISK 479  
Db 163 YVCKLN--ISGIEIASDPLVQLEGPHFIQOEKLNVTNRN--SPFNLTQAV--GP----- 213  
QY 480 KILLYPVNHYEAWQHIQVTNEIVTILNILEPR-----TEVELCVQLVRRGEG-GEHFG 531  
Db 214 ---PEPVEIYFRNNVQLNMK---PYISPSVLVTPGLNETALFSCAHSKGLTASNPG 266  
QY 532 PYRRTTASIGLP--PPRGLNLLPKSQTTMLTWQIPFSSDDFYFVERRRSVQKSDQON 590  
Db 267 QVN-----VKGTPSAKAVHVKRMAHSIVISWVPGF-----DAFSALNSCSVQVKE--- 313  
QY 591 IKVPGNLTSLVLLN-----LHPRQYVVRARVNTKAOCEWSED-----LTAWTLS 635  
Db 314 -AVPOSNSVLLLFNTSVPPHYRQQLWPMADYNI--SVSCNKEYGNSAFSPWITASTTE 370  
QY 636 DILPQPENIKIS-NITHSSAVISWT-----ILDGYSISSITIRYKVOGKNEQHV 685  
Db 371 GAPTQPLNVTVSLNESSFLEIRVWKPLERTHGELOGYH-----IHWTWQDSKGLQNI 425  
QY 686 DVKIK-NATIIQYOLKLEPETAYQVDIFAENNIGS-SNPAP-----SHELVTLPESAQA 739  
Db 456 SLEAQNATVAILPVVATNATCSVRVAATVKGVGFPSPVEVFPASGLITSSPSSTA 485  
QY 740 DLGGKMLLIALILGSGAGTCLTVLLAFLLILO--LKRVNQVRMAQAQFQNVREE----- 791  
Db 486 S-GNTDTSFIVAL-----GFVCGTVAVGLILCLSLVVIQKRCMETKYGNFNRNDSSELV 540  
QY 792 -----PAVQNSGTALNRKKNPNPDTIYPLVDMNDIKPDQVIGENFGQVLKARIK 844  
Db 541 AKKSYCRAREVELTGLSGVSELOQLODV---VIDRNLALSLGKVLGEFGFVMEGRLS 597  
QY 845 K-DGLRMDAAIKRMKEYASKDDHRDFAGEL-EVLC-KLGHHPNINILGACSHRGYLYIA 901  
Db 598 QPEGTPQKVAVKTMK--LDNFSHREIEEFLSEAAQKQFDPNVIKLLGVC-----IELS 650  
QY 902 IEYAP-----HGNLLDFLRKSRVLETPAPAIANSTASTLSQQLLHFAADVARG 951  
Db 651 SQQIPKPMVVLFPFKYGDLSHFLRSR-LEMAQP-----VPLQMLLKFMVYDIALG 700  
QY 952 MDYLSOKOFIHRDLAARNILVGENYVAKIADPGLSR---GOEYVVKTKMRLPVRMAIE 1008  
Db 701 MEYLSRROFLHRDLAARNCLMDMTVCVADPGLSKLYSGDYRQGRIAKNPVMKIAIE 760  
QY 1009 SLNTSVYTTNSDVSYSYGLLWEIVSLGTPYCGMTCAELYEKLPQGYRLEKPLNCDDVY 1068  
Db 761 SLADRVYTTKSDVWAFVGTWMEIATRGMTYPGVQNHIEIYELPHGQRLKKPENCLDELY 820  
QY 1069 DIMRCQWREKPERPSPAQILVSLNRLMEERKT-----YVNTTLYEK 1110  
Db 821 DIMSSCWRAEPADRPFTSQLVHLEKLLLESLEPAPRSGDKDVIYVNTSLPBE 870

RESULT 11  
S38579  
fibroblast growth factor receptor 3 - Iberian ribbed newt (fragment)  
C:Species: Pleurodeles waltlii (Iberian ribbed newt)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
C:Accession: S38579  
R:Shi, D.L.; Fromentoux, V.; Launay, C.; Umbhauer, M.; Boucaut, J.C.  
submitted to the EMBL Data Library, November 1993  
A:Description: Expression of FGFR-3 in amphibian embryos.



A:Reference number: S38579

A:Accession: S38579

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-797 <SH>

A:Cross-references: UNIPROT:Q91287; UNIPARC:UPI0000175635; EMBL:X75603

C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote

C:Keywords: ATP; growth factor receptor

F:258-331/Domain: immunoglobulin homology <IMM>

F:456-741/Domain: protein kinase homology <KIN>

F:464-472/Region: protein kinase ATP-binding motif

Query Match 11.9%; Score 713.5; DB 2; Length 797;

Best Local Similarity 28.7%; Pred. No. 1.4e-21;

Matches 228; Conservative 125; Mismatches 284; Indels 157; Gaps 28;

QY 389 DGTVLHPKDFNTHFSVAITFIHRLPPDSGVVCSVNTVAGWVEKFNISVKVLPKPL 448

DB 71 DGISVDPPTWSHTGQ---KLLKIINVSYDDSGVYSCARQSSEVLR---NVTVRVTDSPS 124

QY 449 NAPNVIDTGHNFVAVINISSEPVFGDPIKSKLLYKVPNVHYEAMQHIQVTHNIVTLNLYE 508

DB 125 SGDDDDDEES-----ESANAFKTPPEWMEKKLLAVP-----AANTV----- 162

QY 509 PRTEYELCVQLVRRGEGEGHGPVRRFTTASIGLPPRGLNLLPKSOTTLNLTWQPIFP 568

DB 163 -----RFRCPAAGKTP-----SITWL----- 179

QY 569 SSEDFFYE-----VERSVOKSDQONTKVP-----GNLTSVLLNNLHP-REQVVRARVNT 619

DB 180 KNGKEFGKEHRIGGIKLRHQQWSLVWSPDRGNVYCVVANKYGTITRETYTLDLVLERT 239

QY 620 K-----AQGEWSEDLTAWTLSDILPQPPENIKISNTHSSAVISWTLIDGYSISITIRYK 675

DB 240 PHRPILQAGFRSNTKVTVVGSDV-----EFHCKVYSDAQPHIQW-----LKHVEVNGS 286

QY 676 VQGNEDQHVQVKIKNATV-----IQYOLKGLEPETAYQVDIFAENNIGSSNPAFSSH 727

DB 287 KFGPDGNFYVTV-LKTAGVNTSDRELEIQF-LRNVTPEDEGEYTCLAGNSIG-----YSH 339

QY 728 E---LVTLPESQAPADLGGGKMLLIIAGSGAGMTCLTVLLAFLIILQKRVANVRMAQA 784

DB 340 HSAMLTVLPPEAPFPDVTDS---VSILAAAGCAVAVILVVIITPYRKMKPSKKTWNTA 395

QY 785 F-----QNVREPAVQFNSGT-LALNRKVKNNPDPTIYPV-----LDW 821

DB 396 TVHKVSKFPLKRQVSELSNSSNNTPLVRITRLSSSDGPMLANVSELELPADPKWELSR 455

QY 822 NDIRFQDVIGEGFQVLKA---RIKQD---GLRMDAAIKRMKEYASKDDHDFAGELEVL 876

DB 456 SRLTLGKPLGEGCFQVVMADAVGIEKKPKNATSVAVKMLKDDATDKLSDLVSEMEM 515

QY 877 KLGHHPNINLLGACEHRGYLYIAEIVAPHGNNLDFLRKSRVLETDPAFAIANSTSL 936

DB 516 KMWGKHNNINLLGACTQDGLPLVLYEYASKNLREYLRRARPPGMDYSFDTCKLPEQL 575

QY 937 SSQQLLHFAADVARGMDVLSQKQFTHRDLAARNLIVGENYVAKTADGLSR---QGEVYV 993

DB 576 TFKOLVSCAYQVARGMEYLSQKCHRDLAARNVLVTDNNVMTKADFGIARDVHNNIDYK 635

QY 994 KKTMRPLVRWMAESLNYSVYTTNSDVMSYGVLLWEIVSLGGTPYCGMTCAELEYEKLPPQ 1053

DB 636 KTTNGRLPVKMAWPAEALFDRVYTHQSDVMSFGVLLWEIFTLGGSPYGPVPEELFKLKE 695

QY 1054 GYRLEKPLNCDDVVDLMQRCWREKPYRPPSFAQLVLSLNMLEERKT---YVNTTL-YEK 1110

DB 696 GHRMDKPAKNCHELYMIMRECWHAVPSQRPFTKQLVEDLDRVLTVTSTDEYLDLSVPPEQ 755

QY 1111 FTYAGID---CSA 1120

DB 756 YSPACPDHSHSCSS 769

RESULT 12

S29840

fibroblast growth factor receptor 1 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004

C:Accession: S29840

R:Yazaki, N.; Fujita, H.; Ohta, M.; Kawasaki, T.; Itoh, N.

Biochim. Biophys. Acta 1172, 37-42, 1993

A:Title: The structure and expression of the FGF receptor-1 mRNA isoforms in rat tissues.

A:Reference number: S29840; MUID:93176824; PMID:8382532

A:Accession: S29840

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-822 <YAZ>

A:Cross-references: UNIPROT:Q04589; UNIPARC:UPI000012A727; EMBL:D12498; NID:Q220737; PID:

C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei

C:Keywords: ATP; autophosphorylation; growth factor receptor; phosphoprotein; transmembr

F:171-232/Domain: immunoglobulin homology <IMM>

F:476-761/Domain: protein kinase homology <KIN>

F:484-492/Region: protein kinase ATP-binding motif

Query Match 11.7%; Score 703.5; DB 2; Length 822;

Best Local Similarity 26.6%; Pred. No. 3.6e-21;

Matches 231; Conservative 141; Mismatches 301; Indels 197; Gaps 28;

QY 334 GWQGL-----QCEREGIPRMTPKIVDLPDHIEVNSGKFNPKICKASGWLPTNEEM 383

DB 3 3GWRGLFWAVLVATLTCTARPAP-----TLPE-----QAQWGVPEVES 42

QY 384 TLVKP-----DGTVLHPKDFNTHFSVAITFIHRLPPDSGVW 422

DB 43 LLVHPGDLQLRLCRDDVQSVINMLRDGVQL--AESNRT-RITGEEVEVRSDIPADSLY 99

QY 423 VCSVNTVAGWVEKFNISVK-VLPKPLNAPNVITDTHGNFAVINISSEPYFGDGPIKSKKL 481

DB 100 ACVTNSPSGSDTTTFYVNVSDALPSEDDDDDDDS-----SSEKETDNTKPNR-- 148

QY 482 LYKPNVHYEAMQHIQVTHNIVTLNLEPRTEYELCVQLVRRGEGEGHGPVRRFTTASI 541

DB 149 --RPVAPY--W-----TSPEKMEKLLHAVPAAKTVKFCPSSTGPTSLRWLKNKG 195

QY 542 GLPPRPGNLNLPKSTOTTLNLTWQPIFPSEDDF--YVEVERRSVOKSDQONI----- 591

DB 196 EFKPDHRIGGKVKYATWSIIMDSVWPSDKNGYTCIVENEYSGINHTYQLDIVERSPHRP 255

QY 592 ----KVPGLTSLVLLNL-----HPREQVVRARVNTKAGQWSEDLTAWTLSDI 637

DB 256 ILQAGLPANKTVALGNSVVEFMCKYSDPQPHIQWLKHIEVNGSKIG----- 301

QY 638 LPPQENIKISNTHSSAVISWTLIDGYSISITIRYKVGKNEHQHVDVKIKNATIIQY 697

DB 302 ----PDNLFPYDQILKTAGV-----NTTDKENEV-----L 326

QY 698 QLKGLEPETAYQVDIFAENNIG-SSNPAFASHLVTLPESQAPADLGGGKMLLIIAILGSAG 756

DB 327 HLRNVSPEDAGEYTCLAGNSIGLSHSAWLTVLEALERPA---VMTSPLYLEIIYCTG 383

QY 757 MTCLTVLLAFLIILQ-----KRVNQVRMA-----QAFQNVREPAVQFNSGTLA 802

DB 384 AFLISCVMGVSIYIKMSGKTKKSDFFHSQMAVHLKLAISPLRRQVTVSADSSASMSNVLL 443

QY 803 LN-RVKKNPNDDPTIYPVLDW---NDIKFO-----DVIGEGNFQVLKARIKKDGLR 849

DB 444 VRPSRLSSGTPMLAGVSEYELPEDRWELPRDLVLGKPLGEGCFGQVFLA--EAILGD 501

QY 850 MD-----AAIKRMKEYASKODHRDFAGELEVLCKLGHHPNIIINLLGACEHRGYLYLAI 902

DB 502 KDKENRVTKVAVKMLKSDATEKLDLISEMEMMKWIGKHKNNIINLLGACTQDGLPLYIV 561

QY 903 EYAPHGNLLDFLRKSRVLETDPAFAIANSTASTTSSQQLHFAADVARGMDVLSQKQTH 962

DB 562 EYASKGNLREYLOARRPPGLECYNPSHNPBEQLSKDLVSCAYQVARGMEYLSKCKCIH 621



QY 963 RDLARNILVGENYAKIADPGLSRG---QEVVVKTKMGLPVRWMAIESLNSVYVTTNS 1019  
Db 622 RDLARNVLVDNWKIADPGLARDIHRIDYIKTKTNGRLPVKQWMAPEALFDRIYTHOS 681  
QY 1020 DWMSGVLLWEIVSLGGTPYCGWTCALYEKLPQGVRLKPLNCDDEVDVLMRCQWREKP 1079  
Db 682 DWMSGVLLWEIVSLGGTPYCGWTCALYEKLPQGVRLKPLNCDDEVDVLMRCQWREKP 741  
QY 1080 YERPSPAQIILVSLNRM--LEERKTYVNTTL 1107  
Db 742 SQRPFTKQLVEDLDRIVALTNSQEYLDLSM 771  
RESULT 13  
JH0393  
N:Contains: growth factor receptor precursor - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jul-1999  
C:Accession: JH0393  
R:Kouhara, H.; Kasayama, S.; Saito, H.; Matsumoto, K.; Sato, B.  
Biochem. Biophys. Res. Commun. 176, 31-37, 1991  
A:Title: Expression cDNA cloning of fibroblast growth factor (FGF) receptor in mouse bre  
A:Reference number: JH0393; MUID:91207411; PMID:1708247  
A:Accession: JH0393  
A:Molecule type: mRNA  
A:Residues: 1-832 <KOU>  
A:Cross-references: UNIPARC:UPI00016CD52; GB:M65053; PIDN:AAA37620.1; PID:  
A:Experimental source: breast cancer cell line SC-3  
A>Note: the authors translated the codon ATG for residue 397 as Ile and CTG for residue  
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote  
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor;  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-832/Product: fibroblast growth factor receptor #status predicted <FGF>  
F:181-242/Domain: immunoglobulin homology <IM>  
F:387-409/Domain: transmembrane #status predicted <TM>  
F:486-771/Domain: protein kinase homology <KIN>  
F:494-502/Region: protein kinase ATP-binding motif  
F:67-113,188-240,287-351/Diulfide bonds: #status predicted  
F:524,541,633/Active site: Lys, Glu, Asp #status predicted  
Query Match 11.7%; Score 701.5; DB 2; Length 832;  
Best Local Similarity 26.9%; Pred. No. 4.3e-21;  
Matches 225; Conservative 133; Mismatches 282; Indels 195; Gaps 25;  
QY 372 ASGWPLPTNEMTLVKPK-----DGTVLHPKDFNHTDHFSAIPT 410  
Db 43 AQPWGVPEVESLLVHPGDLQLRCLRDDVQSIWLRDGVQ--VESNRT-RITGEVE 99  
QY 411 IHRILPPDPSGVWVCSVNTVAGVVEKPFNISVK-VLPKPLNAPNVITDGHFAVINISS- 468  
Db 100 VRDIPADSLYACVTSFSGSDTTTFVSVDALPSSDDDDDDSS-----SSSE 150  
QY 469 -----PYGDDGPIKSKLLYKPNVYEAQHIQVNEIVTLNLYLPRTEYELC 516  
Db 151 KETDNTKPNVAPVYMTSPKMEKKLHAPV-----AKTVFKCP----- 189  
QY 517 VQLVRGEGEGHPGVRFTTASIGLPPRGLNLLPKSQTTTLNLTWQPIFPSSDDF-- 574  
Db 190 -----SSGPNFTLRLWLNKGEKFDHRIGGVKRYATWSIMDSVPSDKGNYTC 240  
QY 575 YVEVERSSVQKSDQNN-----KVPGNLTSLVLLNNL-----HPREQV 612  
Db 241 IVENEYGSINHVTQLDVERSPHRPHPSAGLPANKTVALGNSVFPKCVSDPQPHIQWL 300  
QY 613 VRARVNTKAGWSEDLTAWTLDILPPPENIKISNTHISSAVISWTILDGYSISSITI 672  
Db 301 KHEVNGSKIG-----PDNLPPYVQILKTAGV----- 326  
QY 673 RYKVGKNEQDQVVDVKIKNATIIQYOLKGLBETAQVQDIFANNIG--SSNPAFSHELYT 731  
Db 327 -----NTTDKMEV-----LHLRNVSPFEDAGEYCTCLAGNSIGLSHHSAWLTVLEA 371

QY 732 LPBSQAPADLGGKMLLIALILGASGWTCLTVLLAFILILQL-----KRANVQRMA----- 782  
Db 372 LEERPA---VMTSPYLRIIYCTGAFLMSCLSGSVIIYKMKSGTKGSDPHSQMAVHKLA 428  
QY 783 -----QAFQNVREBPAVQFNSGTLALN-RKVKNNDPDTYIPVLDW---NDIKFQ----- 827  
Db 429 KSIPLRRQVTVSADSSASNSGVLLVRSLSSTGTPMLAGVSEYELPEDPFWELPRDL 488  
QY 828 ---DVIGEGNFQGVILKARIKKDGLRMD-----AAIKRMKEYASKDHRDPAGELEVL 877  
Db 489 VLKPLGEGCGQVLA--EAILGDKDPNRLTKVAVKMLKSDATEKOLSDLISEMMMK 546  
QY 878 KLGHHPNIINILGACHEHGYLYLAIEYAPHGNLLDFLKRSLVLETDPAFANASTASTLS 937  
Db 547 MIGKHNIINILGACTQDGPYVIVEYASKGNLREYLAQRRPPGLECYCNPSPHPEEQUS 606  
QY 938 SOQLHFAADVARGMVLYSOKQFIHRDLAARNILGENYAKIADPGLSRG---QEVVVK 994  
Db 607 SKDLYSCAYQVARGMEYLASKCIHRDLAARNVLTEDNVKMLADFLGLARDIHHIDYKK 666  
QY 995 KTMGLPVRWMAIESLNSVYVTTNSDVMSYGVLLWEIVSLGGTPYCGWTCALYEKLPQ 1054  
Db 667 TTNGRLPVKMAPEALFDRIYTHQSDVMSFGVLLWEIFTLGGSPYGVPEBLFKLLKEG 726  
QY 1055 YRLKPLNCDDEVDVLMRCQWREKPYERPSPAQIILVSLNRM--LEERKTYVNTTL 1107  
Db 727 HRMDKPSNCTNELYMMRDCWHAVPQRPTFKQLVEDLDRIVALTNSQEYLDLSI 781  
RESULT 14  
A48991  
N:Contains: growth factor receptor - mouse  
A:Alternate names: HBGF receptor  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)  
C:Species: Mus musculus (house mouse)  
C>Date: 19-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A48991  
R:Katoch, O.; Hattori, Y.; Sasaki, H.; Sakamoto, H.; Fujimoto, K.; Fujii, T.; Sugimura, T.  
Cancer Res. 53, 1136-1141, 1993  
A:Title: Isolation of the complementary DNA encoding a mouse heparin-binding growth fact  
A:Reference number: A48991; MUID:93177694; PMID:8382556  
A:Accession: A48991  
A>Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-800 <KAT>  
A:Cross-references: UNIPROT:Q61851; UNIPARC:UPI0000170C4F; GB:S56291; NID:G988329; PIDN:  
A:Experimental source: brain  
A>Note: sequence extracted from NCBI backbone (NCBIN:126536, NCBI:126537)  
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei  
C:Keywords: ATP; growth factor receptor; heparin binding; phosphotransferase; tyrosine-s  
F:262-335/Domain: immunoglobulin homology <IM>  
F:464-749/Domain: protein kinase homology <KIN>  
F:472-480/Region: protein kinase ATP-binding motif  
Query Match 11.6%; Score 698.5; DB 2; Length 800;  
Best Local Similarity 28.6%; Pred. No. 5.5e-21;  
Matches 228; Conservative 127; Mismatches 296; Indels 147; Gaps 26;  
QY 388 PDGTVLHPKDFNHTDHFSAIPTIHRIL-----DGTVLASHRILVGPORQLVNASHEDAGVYSQHLRTRVL-- 435  
Db 66 FPGPTVWAKD-----GTGLVASHRILVGPORQLVNASHEDAGVYSQHLRTRVL-- 116  
QY 436 PFNISKVLPKPLNAPNVITDGHFAVINISSPEYFGDGPISKSKLLYKPNVYEAQHI 495  
Db 117 -CHFSVRV----TDAPSSGDDGEDVAEDTCAPYWTRPERMDKKLAVPAANTVFRFCP 171  
QY 496 QVTNEIVTLNLYLEPTTEYELCVQLVRRGEGGHPGVRFTTASIGLPPRGLNLLPKS 555  
Db 172 AAGNPTPSISLWKNKEF-----RGEHRIG--GIKLRHQQSLVME-----SVYPSD 216  
QY 556 QTTNLNTWQPIFPSSDDFYVEVERSSVQKSDQNNIKVPGNLTSLV-----LNNL 605  
Db 217 RGNVTCVVENKFGSIRQTYTLDVLSRPHRPILQ--AGLPANQTALIGSDVEFHCYKVS 275





**this Page Blank (uspto)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.  
OM protein - protein search, using sw model  
Run on: March 14, 2006, 01:40:28 ; Search time 253 Seconds  
(without alignments)  
3134.442 Million cell updates/sec  
Title: US-10-633-742-8  
Perfect score: 6013  
Sequence: 1 MDSLASVLGVSGVLLSGTV.....TTLVKEFTYAGIDCSAEAA 1124

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 2166443 seqs, 705528306 residues  
Total number of hits satisfying chosen parameters: 2166443  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries  
Database : UniProt 05.80.\*  
1: uniprot\_eprot.\*  
2: uniprot\_trembl.\*  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6013	100.0	1124	1 TIE2 HUMAN	Q02763 homo sapien
2	6008	99.9	1157	2 Q59HG2 HUMAN	Q59hg2 homo sapien
3	6000	99.8	1124	2 Q5TCU2 HUMAN	Q5tcu2 homo sapien
4	5981	99.5	1124	2 Q5RDB0 PONPY	Q5rdb0 pongo pygma
5	5964.5	99.2	1123	2 Q5RDI5 PONPY	Q5rdi5 pongo pygma
6	5890	98.0	1101	2 Q8IV34 HUMAN	Q8iv34 homo sapien
7	5732.5	95.3	1125	1 TIE2_BOVIN	Q06807 bos taurus
8	5584	92.9	1122	1 TIE2_MOUSE	Q02858 mus musculus
9	5413	90.0	1083	2 Q9QW24 9MURI	Q9qw24 rattus sp.
10	5296	88.1	1072	2 Q8OYS4 MOUSE	Q8oys4 mus musculus
11	4168.5	69.3	788	2 Q8IXB8 HUMAN	Q8ixb8 homo sapien
12	2915.5	48.5	1116	1 TIE2_BRARE	Q73791 brachydanio
13	2527	42.0	1138	1 TIE1_HUMAN	P35590 homo sapien
14	2526	42.0	1136	1 TIE1_BOVIN	Q06805 bos taurus
15	2521	41.9	1134	2 Q8BG12 MOUSE	Q8bg12 m mus muscu
16	2515	41.8	1134	1 TIE1_MOUSE	Q06806 mus musculus
17	2451	40.8	658	2 Q9OX17 XENLA	Q9ox17 xenopus lae
18	1791	29.8	978	2 Q4RJ67 TETNG	Q4rj67 tetraodon n
19	1524	25.3	421	2 Q9HBS4 HUMAN	Q9hbs4 homo sapien
20	985	16.4	647	2 Q6PAP2 MOUSE	Q6pap2 mus musculus
21	847	14.1	160	2 Q5SH71 SHEEP	Q5sh71 ovie aries
22	839	14.0	184	1 TIE1_BRARE	Q73792 brachydanio
23	719.5	12.0	806	1 CEK2_CHICK	P18460 gallus gall
24	717.5	11.9	974	2 Q90777 CHICK	Q90777 gallus gall
25	715	11.9	816	2 Q86PM4 HYDAT	Q86pm4 hydra atten
26	713.5	11.9	796	2 Q91287 PLEWA	Q91287 pleurodeles
27	712.5	11.8	802	2 Q5SM13 BOVIN	Q5sm13 bos taurus
28	707.5	11.8	683	2 Q5TTJ6 ANOGA	Q5ttj6 anopheles g
29	704	11.7	814	2 Q5SH40 HUMAN	Q5sh40 homo sapien
30	703.5	11.7	822	1 FGFR1_MOUSE	P16092 mus musculus
31	703.5	11.7	822	1 FGFR1_RAT	Q04589 rattus norv

32	702.5	11.7	800	2 Q9JHX9 RAT	Q9jhx9 rattus norv
33	702.5	11.7	820	2 Q8CIM9 MOUSE	Q8cim9 mus musculus
34	699	11.6	802	2 Q42127 XENLA	Q42127 xenopus lae
35	698.5	11.6	800	2 Q7TSI8 MOUSE	Q7tsi8 mus musculus
36	697.5	11.6	820	2 Q5RQ3 PONPY	Q5rq3 pongo pygma
37	697	11.6	729	1 FGFR1_DROME	Q07407 drosophila
38	696.5	11.6	800	2 Q99052 MOUSE	Q99052 mus musculus
39	696.5	11.6	822	2 Q60818 MOUSE	Q60818 mus musculus
40	694.5	11.5	822	1 FGFR1_HUMAN	P11362 homo sapien
41	694	11.5	810	2 Q9PS96 XENLA	Q9ps96 xenopus lae
42	693	11.5	769	2 Q8NI15 HUMAN	Q8ni15 homo sapien
43	693	11.5	806	1 FGFR3_HUMAN	P22607 homo sapien
44	693	11.5	829	2 Q9FSV8 XENLA	Q9psv8 xenopus lae
45	692.5	11.5	828	2 Q9DGK3 XENLA	Q9dgd3 xenopus lae

ALIGNMENTS

RESULT 1  
ID TIE2 HUMAN STANDARD; PRT; 1124 AA.  
AC Q02763;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Angiopoietin 1 receptor precursor (EC 2.7.1.112) (Tyrosine-protein kinase receptor TIE-2) (Tyrosine-protein kinase receptor TEK) (P140 TEK) (Tunica interna endothelial cell kinase) (CD202b antigen).  
DE TEK)  
GN Name=TEK; Synonyms=TIE2;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Placenta;  
RX MEDLINE=93173509; PubMed=8382358;  
RA Ziegler S.F., Bird T.A., Schneringer J.A., Schooley K.A., Baum P.R.;  
RT "Molecular cloning and characterization of a novel receptor protein tyrosine kinase from human placenta."  
RT Oncogene 8:663-670(1993).  
RL [2]  
RP PROTEIN SEQUENCE OF 23-37.  
RX PubMed=15340161; DOI=10.1110/ps.04682504;  
RA Zhang Z., Hensel W.J.;  
RT "Signal peptide prediction based on analysis of experimentally verified cleavage sites."  
RL Protein Sci. 13:2819-2824(2004).  
RN [3]  
RP VARIANT VMCM TRP-849.  
RX MEDLINE=97134665; PubMed=8980225; DOI=10.1016/S0092-8674(00)81814-0;  
RA Vikkula M., Boon L.M., Carraway K.L. III, Calvert J.T., Diamonti A.J., Goumirev B., Pasyk K.A., Marchuk D.A., Warman M.L., Cantley L.C., Mulliken J.B., Olse B.R.;  
RT "Vascular dysmorphogenesis caused by an activating mutation in the receptor tyrosine kinase TIE2."  
RL Cell 87:1181-1190(1996).  
RN [4]  
RP VARIANTS VMCM TRP-849 AND SER-897.  
RX MEDLINE=99299243; PubMed=10369874; DOI=10.1093/hmg/8.7.1279;  
RA Calvert J.T., Riney T.J., Kontos C.D., Cha E.H., Prieto V.G., Shea C.R., Berg J.N., Nevin N.C., Simpson S.A., Pasyk K.A., Speer M.C., Peters K.G., Marchuk D.A.;  
RT "Allelic and locus heterogeneity in inherited venous malformations." Hum. Mol. Genet. 8:1279-1289(1999).  
RL -!  
CC FUNCTION: This protein is a protein tyrosine-kinase transmembrane receptor for angiopoietin 1. It may constitute the earliest mammalian endothelial cell lineage marker. Probably regulates endothelial cell proliferation, differentiation and guides the proper patterning of endothelial cells during blood vessel formation.



SQ SEQUENCE 1124 AA; 125811 MW; 65BC05D18FA4CCEC CRC64;

Query Match 100.0%; Score 6013; DB 1; Length 1124;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSLASLVLCGVSLLSGTVEGAMDLILINSIPLVSDAETSLTCTASGWRPHEPTIGRD 60  
DB 1 MDSLASLVLCGVSLLSGTVEGAMDLILINSIPLVSDAETSLTCTASGWRPHEPTIGRD 60

QY 61 PEALANQODPLEVTDQVTRWAKVWKRKASKINGAYFCGEVRGCAIRIRTMKMQ 120  
DB 61 PEALANQODPLEVTDQVTRWAKVWKRKASKINGAYFCGEVRGCAIRIRTMKMQ 120

QY 121 QASFLPATLTMTVDKGNVNIKFVKLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
DB 121 QASFLPATLTMTVDKGNVNIKFVKLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180

QY 181 AQPQAGVYSARYIGNLTFTSFTRLIVRCEAQKNGPNCNHLCTACMNGVCHEDTGC 240  
DB 181 AQPQAGVYSARYIGNLTFTSFTRLIVRCEAQKNGPNCNHLCTACMNGVCHEDTGC 240

QY 241 ICPPGFMGRTCEKACELHTFGTKERCSCGQGSYVFCPLDPYGCSCATGMKGLQNE 300  
DB 241 ICPPGFMGRTCEKACELHTFGTKERCSCGQGSYVFCPLDPYGCSCATGMKGLQNE 300

QY 301 ACHPGFYGPDCKLRSCNNGEMCDRFGQCLCSFGWQGLQCEREGIPRMTPKIVDLPDHIE 360  
DB 301 ACHPGFYGPDCKLRSCNNGEMCDRFGQCLCSFGWQGLQCEREGIPRMTPKIVDLPDHIE 360

QY 361 VNSGKFNPICKASGWLPTNEMTLVKPDGTVLHPKDFNHTDHFVAFTIHRILPPDSG 420  
DB 361 VNSGKFNPICKASGWLPTNEMTLVKPDGTVLHPKDFNHTDHFVAFTIHRILPPDSG 420

QY 421 VVVCVNTVAGVVEKPNISVVKLPKLNAPNVIDTGNFAVINISSEPYFGDGPIKSKK 480  
DB 421 VVVCVNTVAGVVEKPNISVVKLPKLNAPNVIDTGNFAVINISSEPYFGDGPIKSKK 480

QY 481 LLYKPNVHYEAWQHLQVTNEIVTLNLEPRTEYELCVQLVRGEGEGHPGVPRFTTAS 540  
DB 481 LLYKPNVHYEAWQHLQVTNEIVTLNLEPRTEYELCVQLVRGEGEGHPGVPRFTTAS 540

QY 541 IGLPPRGNLNLPKSGTTLNLTWQPIFSSSEDDFVEVERRSVQKSDQONIKVPGNLTSV 600  
DB 541 IGLPPRGNLNLPKSGTTLNLTWQPIFSSSEDDFVEVERRSVQKSDQONIKVPGNLTSV 600

QY 601 LNNLHPREQYVVRVANTKAQWSEDLTAWTSDILPPQPNIKIKNITHSSAVISWT 660  
DB 601 LNNLHPREQYVVRVANTKAQWSEDLTAWTSDILPPQPNIKIKNITHSSAVISWT 660

QY 661 ILDGYSISITIRYKQGNEDQHVVDVKIKNATIIQYQLKLEPETAYQVDIFAENNIGS 720  
DB 661 ILDGYSISITIRYKQGNEDQHVVDVKIKNATIIQYQLKLEPETAYQVDIFAENNIGS 720

QY 721 SNPAFSELVTLPEAOADLGGKXWLLIAILGSAGMCTCLTVLLAFLIQLKRVANVQR 780  
DB 721 SNPAFSELVTLPEAOADLGGKXWLLIAILGSAGMCTCLTVLLAFLIQLKRVANVQR 780

QY 781 MAQAFONVREEPVAFNSGTALNRKVNPNPTIYPVLWDNDIKFQDVIGEGNGOVLIK 840  
DB 781 MAQAFONVREEPVAFNSGTALNRKVNPNPTIYPVLWDNDIKFQDVIGEGNGOVLIK 840

QY 841 ARIKDDGLRMDAAIKRMKEYASKDHRDPAGBLEVCKLGHHPNIINLIGACEHRGYLYL 900  
DB 841 ARIKDDGLRMDAAIKRMKEYASKDHRDPAGBLEVCKLGHHPNIINLIGACEHRGYLYL 900

QY 901 ALEYAPHGNLDFLRKRSVLETPDPAFANSTASTLSSQQLLHFAADVARGMDYLSQKQF 960  
DB 901 ALEYAPHGNLDFLRKRSVLETPDPAFANSTASTLSSQQLLHFAADVARGMDYLSQKQF 960

QY 961 IHRDLAARNILGENYVAKIADFGLSRGQEVVVKTMGLPVRWMAIESLNVSVYTTNSD 1020  
DB 961 IHRDLAARNILGENYVAKIADFGLSRGQEVVVKTMGLPVRWMAIESLNVSVYTTNSD 1020

QY 1021 VMSYGVLLWEIVSLGSGTFCGWTCAELYEKLPQGYRLEKPLNCDDEVYDLMRQCWRKPY 1080  
DB 1021 VMSYGVLLWEIVSLGSGTFCGWTCAELYEKLPQGYRLEKPLNCDDEVYDLMRQCWRKPY 1080

QY 1081 ERPSAQIILVSLNRMLEERKTYVNTLYEKTYYAGIDCSABEAA 1124  
DB 1081 ERPSAQIILVSLNRMLEERKTYVNTLYEKTYYAGIDCSABEAA 1124

RESULT 2  
Q59HG2 HUMAN PRELIMINARY; PRT; 1157 AA.  
AC Q59HG2;  
DT 10-MAY-2005 (Tremblrel. 30, Created)  
DT 10-MAY-2005 (Tremblrel. 30, Last sequence update)  
DE TEK tyrosine kinase variant (Fragment).  
GN Name=TEK tyrosine kinase variant;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Brain;  
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,  
RA Ohara O., Nagase T., Kikuno F.R.;  
RT "None Title";  
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB208796; BAD2033.1; -; mRNA.  
DR SMR; Q59HG2; 846-1154.  
DR Ensembl; ENSG00000120156; Homo sapiens.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0008201; F:heparin binding; IEA.  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR000742; EGF\_2.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR003961; FN\_III.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR002290; Ser\_thr\_kinase.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR InterPro; IPR008266; Tyr\_kinase\_AS.  
DR Pfam; PF00041; fn3; 3.  
DR PRINTS; PR00109; TYRKINASE.  
DR SMART; SM00181; EGF; 3.  
DR SMART; SM00060; FN3; 3.  
DR SMART; SM00220; S\_TK; 1.  
DR SMART; SM00219; TyrK; 1.  
DR PROSITE; PS00022; EGF\_1; 3.  
DR PROSITE; PS01186; EGF\_2; 3.  
DR PROSITE; PS50026; EGF\_3; 1.  
DR PROSITE; PS50853; FN3; 3.  
DR PROSITE; PS50835; IG\_LIKE; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
KW ATP-binding; EGF-like domain; Immunoglobulin domain; Kinase; Receptor;  
KW Transmembrane.  
FT NON\_TER 1  
SQ SEQUENCE 1157 AA; 129275 MW; 872BFE11E6CDCF56 CRC64;

Query Match 99.9%; Score 6008; DB 2; Length 1157;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1123; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MDLASLVLGVSLLSGTVEGAMDLILINSPLVSDAETSLTCIASGWRPHEPTIGRD 60
DB 34 MDLASLVLGVSLLSGTVEGAMDLILINSPLVSDAETSLTCIASGWRPHEPTIGRD 93
QY 61 FEALMNQHDPLEVTDVTTREWAKVWVKREKASKINGAYFCEGVRGAEIRITMKVRQ 120
DB 94 FEALMNQHDPLEVTDVTTREWAKVWVKREKASKINGAYFCEGVRGAEIRITMKVRQ 153
QY 121 QASFLPATLTMTVDKGNVTSFKKVLKEEDAVIYKNGSFTHSVPRHEVDILEVHLPH 180
DB 154 QASFLPATLTMTVDKGNVTSFKKVLKEEDAVIYKNGSFTHSVPRHEVDILEVHLPH 213
QY 181 AQPQDAGVYSARYIGCNLFSAFTRLIVRRCEAQKWGPCNHLCTACMNGVCHDTEGEC 240
DB 214 AQPQDAGVYSARYIGCNLFSAFTRLIVRRCEAQKWGPCNHLCTACMNGVCHDTEGEC 273
QY 241 ICPPGFMGRTCEKACELHTFGRTCKERCSGQGGCKSYVFCPLPDPYGCSCATGWKGLQNE 300
DB 274 ICPPGFMGRTCEKACELHTFGRTCKERCSGQGGCKSYVFCPLPDPYGCSCATGWKGLQNE 333
QY 301 ACHPGFYGPDCKLRCSNNGECMDRFOGCLSPGWQGLQCEBREGIPRMTPKIVLDPDHE 360
DB 334 ACHPGFYGPDCKLRCSNNGECMDRFOGCLSPGWQGLQCEBREGIPRMTPKIVLDPDHE 393
QY 361 VNSGKFNPICKASGWPLPTNEBMTLVKPDGTVLHPKDFNHTDHFSAFTTHRIILPPDSG 420
DB 394 VNSGKFNPICKASGWPLPTNEBMTLVKPDGTVLHPKDFNHTDHFSAFTTHRIILPPDSG 453
QY 421 VVWCSVNTVAGVKEKPNISVKVLKPKLNAPNVIDTGNFVAINISSPYFGDGPIKSKK 480
DB 454 VVWCSVNTVAGVKEKPNISVKVLKPKLNAPNVIDTGNFVAINISSPYFGDGPIKSKK 513
QY 481 LLYKPVNHYEAWQHIVQNTNEIVTLNLYLPRTEYELCVQLVRRGEGCHPGVRRFTTAS 540
DB 514 LLYKPVNHYEAWQHIVQNTNEIVTLNLYLPRTEYELCVQLVRRGEGCHPGVRRFTTAS 573
QY 541 IGLPPRGNLNLPKSOITLNTWQPIFSSSEDDFVVEVERSVQKSDOONIKVPGNLTSV 600
DB 574 IGLPPRGNLNLPKSOITLNTWQPIFSSSEDDFVVEVERSVQKSDOONIKVPGNLTSV 633
QY 601 LLNNLHPREQVYVRVAVNTKQAGSEDLTAWTILSDILPPDENIKISNITHSSAVISWT 660
DB 634 LLNNLHPREQVYVRVAVNTKQAGSEDLTAWTILSDILPPDENIKISNITHSSAVISWT 693
QY 661 ILDGYSISSITIRYKQGNEDQHVVDVKIKNATIIQYQLKLEPETAYOVDFIAENNIGS 720
DB 694 ILDGYSISSITIRYKQGNEDQHVVDVKIKNATIIQYQLKLEPETAYOVDFIAENNIGS 753
QY 721 SNPAFSHELVTLPESQAPADLGGGKMLLIIAIIIGSAGMTCLTVLLAFLIILQIKRANVQR 780
DB 754 SNPAFSHELVTLPESQAPADLGGGKMLLIIAIIIGSAGMTCLTVLLAFLIILQIKRANVQR 813
QY 781 MAQAFQNVREBPVQFNSGTALNRKVRKNPDPTIYPVLWNNDIKFQDVI GEGNFGQVLK 840
DB 814 MAQAFQNVREBPVQFNSGTALNRKVRKNPDPTIYPVLWNNDIKFQDVI GEGNFGQVLK 873
QY 841 ARIKQGLRMDAAIKRMKEYASKDHRDFAGELEVLCVKGHPNIIINLLGACEHRYLYL 900
DB 874 ARIKQGLRMDAAIKRMKEYASKDHRDFAGELEVLCVKGHPNIIINLLGACEHRYLYL 933
QY 901 AIEYAPHGNLDDFLKRSVLETDPAFAIANSTASTLSSQQLLHFAADVARGMVDYLSQKF 960
DB 934 AIEYAPHGNLDDFLKRSVLETDPAFAIANSTASTLSSQQLLHFAADVARGMVDYLSQKF 993
QY 961 IHRDLAARNILVGENYAKIADFGLSRGQEVVYKKTGMRLPVRRMAIESLNYSVYTTNSD 1020
DB 994 IHRDLAARNILVGENYAKIADFGLSRGQEVVYKKTGMRLPVRRMAIESLNYSVYTTNSD 1053
QY 1021 VWSYGVLLWEIVSLGGTFCGWTCAELYEKUPQGYRLEKPLNCDDDEVYDLMRQCWREKPY 1080
DB 1054 VWSYGVLLWEIVSLGGTFCGWTCAELYEKUPQGYRLEKPLNCDDDEVYDLMRQCWREKPY 1113
QY 1081 ERPSFAQLVSLNRMLEERKTYVNTTLYEKTYAGIDCSABEAA 1124
```

```
DB 1114 ERPSFAQLVSLNRMLEERKTYVNTTLYEKTYAGIDCSABEAA 1157
RESULT 3
Q5TCU2 HUMAN
ID Q5TCU2_HUMAN PRELIMINARY; PRT; 1124 AA.
AC Q5TCU2_
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE TEK tyrosine kinase, endothelial (Venous malformations, multiple
DE cutaneous and mucosal).
GN Name=TEK; ORFNames=RP11-57P14.1-002;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Corby N.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL133411; CA116055.1; -; Genomic DNA.
DR EMBL; AL355432; CA116055.1; JOINED; Genomic DNA.
DR EMBL; AL355433; CA116055.1; JOINED; Genomic DNA.
DR SMR; Q5TCU2; 813-1121.
DR Ensemble; ENSG00000120156; Homo sapiens.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00041; fn3; 3.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00181; EGF_3; 3.
DR SMART; SM00080; FN3; 3.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS00026; EGF_3; 1.
DR PROSITE; PS0853; FN3; 3.
DR PROSITE; PS0835; IG LIKE; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Kinase.
SQ SEQUENCE 1124 AA; 125830 MW; E739DEC3E4FE124 CRC64;
Query Match 99.8%; Score 6000; DB 2; length 1124;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1122; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MDLASLVLGVSLLSGTVEGAMDLILINSPLVSDAETSLTCIASGWRPHEPTIGRD 60
DB 1 MDLASLVLGVSLLSGTVEGAMDLILINSPLVSDAETSLTCIASGWRPHEPTIGRD 60
QY 61 FEALMNQHDPLEVTDVTTREWAKVWVKREKASKINGAYFCEGVRGAEIRITMKVRQ 120
DB 61 FEALMNQHDPLEVTDVTTREWAKVWVKREKASKINGAYFCEGVRGAEIRITMKVRQ 120
QY 121 QASFLPATLTMTVDKGNVTSFKKVLKEEDAVIYKNGSFTHSVPRHEVDILEVHLPH 180
```



Db 121 QASFLPATLTMTVDKGDVNIISFKKVLKEBDVYKNGSFHSPRHEVPDILEVHLPH 180  
QY 181 AQPDAGVYSARYIGNLFTSAFTRLIVRCEAQKMGPECNHLCTACMNGVCHEDTGE 240  
Db 181 AQPDAGVYSARYIGNLFTSAFTRLIVRCEAQKMGPECNHLCTACMNGVCHEDTGE 240  
QY 241 ICPPGFMGTCEKACELHTFGTCKERCSDGCKSYVFCPLDPDYGCCATCWKGLQNE 300  
Db 241 ICPPGFMGTCEKACELHTFGTCKERCSDGCKSYVFCPLDPDYGCCATCWKGLQNE 300  
QY 301 ACHPGFYGDCCKLRCSNNGECDFOGCLCSFGWGLQCEGREGIPRTMPKIVLDPDHE 360  
Db 301 ACHPGFYGDCCKLRCSNNGECDFOGCLCSFGWGLQCEGREGIPRTMPKIVLDPDHE 360  
QY 361 VNSGKFNPICKASGWLPTNEEMTLVKPDGTGVLHPKDNFHTDHFSAIFTIHRILPPDSG 420  
Db 361 VNSGKFNPICKASGWLPTNEEMTLVKPDGTGVLHPKDNFHTDHFSAIFTIHRILPPDSG 420  
QY 421 VVVCVNTVAGVVEKPFNISKVLPKPLNAPNVIDTGHFAVINISSEPFYFGDGPIKSK 480  
Db 421 VVVCVNTVAGVVEKPFNISKVLPKPLNAPNVIDTGHFAVINISSEPFYFGDGPIKSK 480  
QY 481 LLYKPNHYEAMQHILQVTNEIYTLNLYLEPTEYELCVQLVRRGEGEGHGPVRRPTTAS 540  
Db 481 LLYKPNHYEAMQHILQVTNEIYTLNLYLEPTEYELCVQLVRRGEGEGHGPVRRPTTAS 540  
QY 541 IGLPPRGNLNLPKSGQTLNLTWQPIFPSSSEDDFYVEVERRSVQSDQONIKVPGNLTSV 600  
Db 541 IGLPPRGNLNLPKSGQTLNLTWQPIFPSSSEDDFYVEVERRSVQSDQONIKVPGNLTSV 600  
QY 601 LNNLHPREYVVRARVNTKAGWSEDLTAWLSDILPPQENIKSNIHSSAVISWT 660  
Db 601 LNNLHPREYVVRARVNTKAGWSEDLTAWLSDILPPQENIKSNIHSSAVISWT 660  
QY 661 ILDGYSISITIRYKVGKNEQHVVDKIKNATIIOYQLKLEPETAYOVDFIAENNIGS 720  
Db 661 ILDGYSISITIRYKVGKNEQHVVDKIKNATIIOYQLKLEPETAYOVDFIAENNIGS 720  
QY 721 SNPAFSHELVTLPESQAPADLGGGKMLLIALIGSAGMTCLTVLLAFLIILQKRANVQR 780  
Db 721 SNPAFSHELVTLPESQAPADLGGGKMLLIALIGSAGMTCLTVLLAFLIILQKRANVQR 780  
QY 781 MAQAFQNVREPAVFNSTGLALNRKVNPNPTIYVLDWMDIKFQDVIGBGNFGVLK 840  
Db 781 MAQAFQNVREPAVFNSTGLALNRKVNPNPTIYVLDWMDIKFQDVIGBGNFGVLK 840  
QY 841 ARIKDGRLMDAAIKRMKEYASKDHRDPAGLEVLCKLGHHPNIIINLLGACEHRYLYL 900  
Db 841 ARIKDGRLMDAAIKRMKEYASKDHRDPAGLEVLCKLGHHPNIIINLLGACEHRYLYL 900  
QY 901 ATEYAPHGNLLDPLKRSVLETDPAFANSTASTLSSQQLLHFAADVARGMDYLSQKF 960  
Db 901 ATEYAPHGNLLDPLKRSVLETDPAFANSTASTLSSQQLLHFAADVARGMDYLSQKF 960  
QY 961 IHRDLAARNILVGENYVAKIADFGLSRGQEVVYKKTGMRLPVRRMAISLNYSVYTTNSD 1020  
Db 961 IHRDLAARNILVGENYVAKIADFGLSRGQEVVYKKTGMRLPVRRMAISLNYSVYTTNSD 1020  
QY 1021 VWSYGVLLWEIVSLGGTPYCGMTCAELYKLPQGYRLEKPLNCDDEVYDLMRQCREKPY 1080  
Db 1021 VWSYGVLLWEIVSLGGTPYCGMTCAELYKLPQGYRLEKPLNCDDEVYDLMRQCREKPY 1080  
QY 1081 ERPSFAQLVLSNRMLEERKTVNTTLYEKFTYAGIDCSAEBA 1124  
Db 1081 ERPSFAQLVLSNRMLEERKTVNTTLYEKFTYAGIDCSAEBA 1124

## RESULT 4

QSRDBO\_PONPY

ID QSRDBO\_PONPY PRELIMINARY; PRT; 1124 AA.

AC QSRDBO\_

DT 01-FEB-2005 (TREMBLrel. 29, Created)

DT 01-FEB-2005 (TREMBLrel. 29, Last sequence update)

DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)  
DE Hypothetical protein DKFPz469L125.  
GN Name=DKFPz469L125;  
OS Pongo pygmaeus (Orangutan).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Pongo.  
OX NCBI\_TaxID=9600;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Kidney;  
RG The German cDNA Consortium;  
RA Ansoerge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,  
RL Newes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;  
Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; CR858004; CAH90247.1; -; mRNA.  
DR SMR; QSRDBO; 813-1121.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR000742; EGF 2.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR003961; FN III.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR000719; Prot kinase.  
DR InterPro; IPR002290; Ser Thr\_kinase.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR InterPro; IPR008266; Tyr\_kinase\_AS.  
DR Pfam; PF00041; fn3; 3.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot kinase; 1.  
DR SMART; SM00181; EGF; 3.  
DR SMART; SM00060; FN3; 3.  
DR SMART; SM00220; S\_TKc; 1.  
DR SMART; SM00219; TyrKc; 1.  
DR PROSITE; PS00022; EGF\_1; 3.  
DR PROSITE; PS01186; EGF\_2; 3.  
DR PROSITE; PS00026; EGF\_3; 1.  
DR PROSITE; PS00853; FN3; 3.  
DR PROSITE; PS00835; IG-LIKE; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 1124 AA; 125926 MW; A771C45B994F09AA2 CRC64;

Query Match 99.5%; Score 5981; DB 2; Length 1124;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1117; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
QY 1 MDSLASLVLCGVSLLSGTVGAMDLILINSLPLVSDAETSLTCTASGRWHPETIGRD 60  
Db 1 MDSLASLVLCGVSLPSGTVGAMDLILINSLPLVSDAETSLTCTASGRWHPETIGRD 60  
QY 61 FEALMNQHDPLEVTQDVTRWAKVWVKREKASKINGAYFCEGVRGGEAIRITWKRQ 120  
Db 61 FEALMNQHDPLEVTQDVTRWAKVWVKREKASKINGAYFCEGVRGGEAIRITWKRQ 120  
QY 121 QASFLPATLTMTVDKGDVNIISFKKVLKEBDVYKNGSFHSPRHEVPDILEVHLPH 180  
Db 121 QASFLPATLTMTVDKGDVNIISFKKVLKEBDVYKNGSFHSPRHEVPDILEVHLPH 180  
QY 181 AQPDAGVYSARYIGNLFTSAFTRLIVRCEAQKMGPECNHLCTACMNGVCHEDTGE 240  
Db 181 AQPDAGVYSARYIGNLFTSAFTRLIVRCEAQKMGPECNHLCTACMNGVCHEDTGE 240  
QY 241 ICPPGFMGTCEKACELHTFGTCKERCSDGCKSYVFCPLDPDYGCCATCWKGLQNE 300  
Db 241 ICPPGFMGTCEKACELHTFGTCKERCSDGCKSYVFCPLDPDYGCCATCWKGLQNE 300



Db 421 VVWCVNTVAGVMEKPFNISVKVLKPKLAPNPNVDTGHNFAVNTSSSEPYFGDGPISKXK 480  
Qy 481 LLYKPNVHYEAWQHIVQVNEIVTLNLEPRTEYELCVQLVRREGEHGPVRFRTTAS 540  
Db 481 LLYKPNVHYEAWQHIVQVNEIVTLNLEPRTEYELCVQLVRREGEHGPVRFRTTAS 540  
Qy 541 IGLPPRGVNLPLPKSQTLLNLTWQIPFSSSEDDFVVEVRRSVQKSDQONIKVPGNLTSV 600  
Db 541 IGLPPRGVNLPLPKSQTLLNLTWQIPFSSSEDDFVVEVRRSVQKSDQONIKVPGNLTSV 600  
Qy 601 LNNLHPREQVYVVRVNTKAGSESDLTATLSDILPPQENIKSNITHSSAVISWT 660  
Db 601 LNNLHPREQVYVVRVNTKAGSESDLTATLSDILPPQENIKSNITHSSAVISWT 660  
Qy 661 ILDGYSISITIRYKVGKNEQHDVVKIKNATIIQYQLKLEPETAYQVDIFAENNIGS 720  
Db 661 ILDGYSISITIRYKVGKNEQHDVVKIKNATIIQYQLKLEPETAYQVDIFAENNIGS 720  
Qy 721 SNPAFSHELVTLPESQAPADLGGGRKMLLAIILGSAGMTCLTVLLAFLIILQLKRANVQR 780  
Db 721 SNPAFSHELVTLPESQAPADLGGGRKMLLAIILGSAGMTCLTVLLAFLIILQLKRANVQR 780  
Qy 781 MAQAFQNVREPAVQNSGTLALNRKVNPPDTIYPVLDMNDIKFQDVIGSGNFGQVLK 840  
Db 781 MAQAFQNVREPAVQNSGTLALNRKVNPPDTIYPVLDMNDIKFQDVIGSGNFGQVLK 840  
Qy 841 ARIKXGLRMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNTIINLIGACEHRGYLYL 900  
Db 840 ARIKXGLRMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNTIINLIGACEHRGYLYL 900  
Qy 901 AIEYAPHGNLLDFLKRSLVETDPAPAFANSTASTLSSQQLLHFAADVARGMDYLSQKQF 960  
Db 900 AIEYAPHGNLLDFLKRSLVETDPAPAFANSTASTLSSQQLLHFAADVARGMDYLSQKQF 960  
Qy 961 THRDLAARNILGENYVAKIADFGLSRGOEVVYKTMGRLPVRWMAIESLNVSVYTTNSD 1020  
Db 960 THRDLAARNILGENYVAKIADFGLSRGOEVVYKTMGRLPVRWMAIESLNVSVYTTNSD 1019  
Qy 1021 VWSYGVLLWEIVSLGTPCYCGMTCAELYEKLPQGRLEKPLNCDDEVYDLMRQWREKPY 1080  
Db 1020 VWSYGVLLWEIVSLGTPCYCGMTCAELYEKLPQGRLEKPLNCDDEVYDLMRQWREKPY 1079  
Qy 1081 ERPSAQILVSLNRLMERKTYVNTLYEKFYTAGIDCSAEAA 1124  
Db 1080 ERPSAQILVSLNRLMERKTYVNTLYEKFYTAGIDCSAEAA 1123

RESULT 6  
Q8IV34 HUMAN PRELIMINARY; PRT; 1101 AA.  
AC Q8IV34;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE TEK protein.  
GN Name=TEK;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Pancreas;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udgin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Pancreas;  
RG NIH MGC Project;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBSJ databases.  
DR EMBL; BC035514; AAH35514.1; -, mRNA.  
DR HSSP; Q02763; IFVR.  
DR SMR; Q02763; 790-1098.  
DR Ensembl; ENSG00000120156; Homo sapiens.  
DR GO; GO:000524; F:ATP binding; IEA.  
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR00742; EGF\_2.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR003961; FN III.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR InterPro; IPR008266; Tyr\_pkinase\_AS.  
DR Pfam; PF00041; fn3; 3.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00181; EGF; 2.  
DR SMART; SM00060; FN3; 3.  
DR SMART; SM00219; TyrKc; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_3.  
DR PROSITE; PS01186; EGF\_2; 3.  
DR PROSITE; PS00026; EGF\_3; 1.  
DR PROSITE; PS00853; FN3; 3.  
DR PROSITE; PS00835; IG LIKE; 1.  
DR PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN\_1.  
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
SQ SEQUENCE 1101 AA; 12359 MW; C9DF2721F767B0D0 CRC64;

Query Match 98.0%; Score 5890; DB 2; Length 1101;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1098; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 24 MDLILNLSPLVSDAETSLTCTIACSWRPHEPTITIGDFEALMNQHDPLEVTDVTRWA 83  
Db 1 MDLILNLSPLVSDAETSLTCTIACSWRPHEPTITIGDFEALMNQHDPLEVTDVTRWA 60  
Qy 84 KKVWVKREKASKINGAYFCGRVGEAIRIRTMKMRQASFLPATLTMVTKGDNVNI 143  
Db 61 KKVWVKREKASKINGAYFCGRVGEAIRIRTMKMRQASFLPATLTMVTKGDNVNI 120  
Qy 144 KKVLIKEEDAVIYKNGSFIHSPVREHVPDILLEVHLPHAQPDAGVYSARYIGNLT 203  
Db 121 KKVLIKEEDAVIYKNGSFIHSPVREHVPDILLEVHLPHAQPDAGVYSARYIGNLT 180  
Qy 204 TRLIVRCEAKWGPCNHLCTACNNNGVCHEDTGECICPPGFMGRTCEKACELHTFG 263  
Db 181 TRLIVRCEAKWGPCNHLCTACNNNGVCHEDTGECICPPGFMGRTCEKACELHTFG 240  
Qy 264 CKERCSGEGGCKSVYFCLPDPYVGCSCATGKGLQCNACHPGFGPDCKLRCSNNNG 323  
Db 241 CKERCSGEGGCKSVYFCLPDPYVGCSCATGKGLQCNACHPGFGPDCKLRCSNNNG 300  
Qy 324 DRFOGCLCSPGWQGLQCRREGIPRMTPKIVLDPDHIENVSGKFNPKICKASGWPL 383

```
Db 301 DRFGCLCPGQWGLQCEGREGIPRTPKIVLDPDHIENVSGKFNFIKASGWPLFTNEEM 360
Qy 384 TLVRKPDGTVLHPKDFNHDHDSVAIFTHRLPPDSGVWVCSVNTVAGWKEPFIISVKV 443
Db 361 TLVRKPDGTVLHPKDFNHDHDSVAIFTHRLPPDSGVWVCSVNTVAGWKEPFIISVKV 420
Qy 444 LPKPLNAPNVDITGHNFAVINISBPYPGDPGPIKSKLLYKPVNHYEAWQHIOVTNEIVT 503
Db 421 LPKPLNAPNVDITGHNFAVINISBPYPGDPGPIKSKLLYKPVNHYEAWQHIOVTNEIVT 480
Qy 504 LNYLPRTEYELCVQLVRGEGEGHPGVRRFTTASIGLPPRGLNLLPKSQTTLNLTW 563
Db 481 LNYLPRTEYELCVQLVRGEGEGHPGVRRFTTASIGLPPRGLNLLPKSQTTLNLTW 540
Qy 564 QPIPFSSDDDFVEVERRSVQKSDQONIKVPGNLTSLVLLNHLHPREQVVRVAVNTKAQG 623
Db 541 QPIPFSSDDDFVEVERRSVQKSDQONIKVPGNLTSLVLLNHLHPREQVVRVAVNTKAQG 600
Qy 624 EWSDELTAWTLSLILPPDENIKISNITHSSAVISWTILDGYSISSIIRYKVGQKNEDQ 683
Db 601 EWSDELTAWTLSLILPPDENIKISNITHSSAVISWTILDGYSISSIIRYKVGQKNEDQ 660
Qy 684 HVDVKIKNATTIQYOLKGLEPETAQYVDI FAENNIGSSNPAPFSHELVTLPESQAPADLGG 743
Db 661 HVDVKIKNATTIQYOLKGLEPETAQYVDI FAENNIGSSNPAPFSHELVTLPESQAPADLGG 720
Qy 744 GKMLLIATLGSAGMTCLTVLLAFLIILQIKRANVQRMAQAFQNVREPAQVFNSTGLAL 803
Db 721 GKMLLIATLGSAGMTCLTVLLAFLIILQIKRANVQRMAQAFQNVREPAQVFNSTGLAL 780
Qy 804 NRKVKNNPDPTIYPVLWDNDIKFDQVIGEGNFGQVLKARIKDGRLMDAAIKRMKEYASK 863
Db 781 NRKVKNNPDPTIYPVLWDNDIKFDQVIGEGNFGQVLKARIKDGRLMDAAIKRMKEYASK 840
Qy 864 DHRDFAGELEVLCKLGHHPNIINLGCAGEHRGYLYLAIEYAPHGNLLDFLKRSLVLETD 923
Db 841 DHRDFAGELEVLCKLGHHPNIINLGCAGEHRGYLYLAIEYAPHGNLLDFLKRSLVLETD 900
Qy 924 PAFALANSTASTLSQQILHFAADVARDGMDYLSQKQFTHRLAARNILVGENYAKIADP 983
Db 901 PAFALANSTASTLSHHLLHFAADVARDGMDYLSQKQFTHRLAARNILVGENYAKIADP 960
Qy 984 GLSRQGVYVKTWGRLPVRWMAIESLNSVYVTTNSDVMSVGLLWEIVSLGGTPYCGMT 1043
Db 961 GLSRQGVYVKTWGRLPVRWMAIESLNSVYVTTNSDVMSVGLLWEIVSLGGTPYCGMT 1020
Qy 1044 CAELYEKLPGQYRLEKPLNCDDDEVYDLMRQCWREKPYERPSFAQILVSLNRLMEERKTYV 1103
Db 1021 CAELYEKLPGQYRLEKPLNCDDDEVYDLMRQCWREKPYERPSFAQILVSLNRLMEERKTYV 1080
Qy 1104 NTTYLYEKFTYAGIDCSAEBA 1124
Db 1081 NTTYLYEKFTYAGIDCSAEBA 1101
```

## RESULT 7

```
TIE2_BOVIN STANDARD; PRT; 1125 AA.
ID TIE2_BOVIN
AC Q06807;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Angiopoietin 1 receptor precursor (EC 2.7.1.112) (Tyrosine-protein
DE kinase receptor TIE-2).
GN Name=TEK; Synonyms=TIE-2, TIE2;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
```

```
RC TISSUE=Endothelial cell;
RX MEDLINE=94022374; PubMed=8415706;
RA Sato T.N., Qin Y., Kozak C.A., Andus K.L.;
RT "Tie-1 and Tie-2 define another class of putative receptor tyrosine
RT kinase genes expressed in early embryonic vascular system.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:9355-9358(1993).
CC -!- FUNCTION: This protein is a protein tyrosine-kinase transmembrane
CC receptor for angiotensin II. It may constitute the earliest
CC mammalian endothelial cell lineage marker. Probably regulates
CC endothelial cell proliferation, differentiation and guides the
CC proper patterning of endothelial cells during blood vessel
CC formation.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + a protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Specifically expressed in developing vascular
CC endothelial cells.
CC -!- SIMILARITY: Belongs to the Tyr protein kinase family. Tie
CC subfamily.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 3 fibronectin type-III domains.
CC -!- SIMILARITY: Contains 2 Ig-like C2-type (immunoglobulin-like)
CC domains.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; X71424; CAA50555.1; -; mRNA.
DR PIR; S57846; S57846.
DR HSSP; Q02763; 1FVR.
DR SMR; Q06807; 814-1122.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR Pfam; PF00041; fn3; 3.
DR PRINTS; PR0109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00180; EGF_Lam; 1.
DR SMART; SM00060; FN3; 3.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00026; EGF_3; 1.
DR PROSITE; PS08553; FN3; 3.
DR PROSITE; PS08335; IG_LIKE; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; EGF-like domain; Glycoprotein; Immunoglobulin domain;
KW Kinase; Multigene family; Nucleotide-binding; Phosphorylation;
KW Receptor; Repeat; Signal; Transmembrane; Tyrosine-protein kinase.
KW Tyrosine-protein kinase.
FT SIGNAL 1 18
FT CHAIN 19 1125 Angiopoietin 1 receptor.
FT TOPO_DOM 19 746 Extracellular (Potential).
FT TRANSMEM 747 771 Potential.
FT TOPO_DOM 772 1125 Cytoplasmic (Potential).
FT DOMAIN 44 123 Ig-like C2-type 1.
FT DOMAIN 210 252 EGF-like 1.
FT DOMAIN 254 299 EGF-like 2.
FT DOMAIN 301 341 EGF-like 3.
FT DOMAIN 350 440 Ig-like C2-type 2.
FT DOMAIN 445 537 Fibronectin type-III 1.
FT DOMAIN 544 634 Fibronectin type-III 2.
FT DOMAIN 640 731 Fibronectin type-III 3.
```



RP NUCLEOTIDE SEQUENCE, AND TISSUE SPECIFICITY.  
 RC TISSUE-Hematopoietic stem cells;  
 RX PubMed-8395828;  
 RA Iwama A., Hamaguchi I., Hashiyama M., Murayama Y., Yasunaga K.,  
 RA Suda T.;  
 RT "Molecular cloning and characterization of mouse TIE and TEK receptor  
 RT tyrosine kinase genes and their expression in hematopoietic stem  
 RT cells.";  
 RL Biochem. Biophys. Res. Commun. 195:301-309(1993).  
 RN [7]  
 RP NUCLEOTIDE SEQUENCE OF 822-1122.  
 RC STRAIN=CD-1; TISSUE=Embryonic heart;  
 RX MEDLINE=923334855; PubMed=1630810;  
 RA Dumont D.J., Yamaguchi T.P., Conlon R.A., Rossant J., Breitman M.L.;  
 RT "Tek, a novel tyrosine kinase gene located on mouse chromosome 4, is  
 RT expressed in endothelial cells and their presumptive precursors.";  
 RL Oncogene 7:1471-1480(1992).  
 CC -!- FUNCTION: This protein is a protein tyrosine-kinase transmembrane  
 CC receptor for angiotensin II. It may constitute the earliest  
 CC mammalian endothelial cell lineage marker. Probably regulates  
 CC endothelial cell proliferation, differentiation and guides the  
 CC proper patterning of endothelial cells during blood vessel  
 CC formation.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + a protein  
 CC tyrosine phosphate.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: Specifically expressed in developing vascular  
 CC endothelial cells. Abundantly expressed in lung and heart,  
 CC moderately in brain, liver and kidney, and weakly in thymus,  
 CC spleen and testis.  
 CC -!- DEVELOPMENTAL STAGE: Expression detectable in day 8.5 embryos.  
 CC -!- SIMILARITY: Belongs to the Tyr protein kinase family. Tie  
 CC subfamily.  
 CC -!- SIMILARITY: Contains 3 EGF-like domains.  
 CC -!- SIMILARITY: Contains 3 fibronectin type-III domains.  
 CC -!- SIMILARITY: Contains 2 Ig-like C2-type (immunoglobulin-like)  
 CC domains.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 DR EMBL; X71426; CAA50857.1; -; mRNA.  
 DR EMBL; X67553; CAA47857.1; -; mRNA.  
 DR EMBL; D13738; BAA02883.1; -; mRNA.  
 DR EMBL; S67051; AAB28663.1; -; mRNA.  
 DR PIR; JN0712; JN0712.  
 DR HSP; Q02763; 1FVR.  
 DR SMR; Q02858; 811-1119.  
 DR Ensembl; ENSMUSG00000006386; Mus musculus.  
 DR MGI; MGI:98664; Tek.  
 DR GO; GO:0005615; C:extracellular space; TAS.  
 DR GO; GO:0016021; C:integral to membrane; TAS.  
 DR GO; GO:0016337; P:cell-cell adhesion; IMP.  
 DR GO; GO:0007160; P:cell-matrix adhesion; IMP.  
 DR GO; GO:0045765; P:regulation of angiogenesis; IMP.  
 DR GO; GO:0030334; P:regulation of cell migration; IMP.  
 DR InterPro; IPR000742; EGF 2.  
 DR InterPro; IPR006209; EGF-like.  
 DR InterPro; IPR003961; FN III.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR002049; Laminin EGF.  
 DR InterPro; IPR000719; Prot kinase.  
 DR InterPro; IPR001245; Tyr pkinase.  
 DR InterPro; IPR008266; Tyr\_pkinase\_AS.  
 DR Pfam; PF00041; fn3; 3.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD0000001; Prot kinase; 1.  
 DR SMART; SM00180; EGF\_Lam; 1.  
 DR SMART; SM00060; FN3; 3.  
 DR SMART; SM00219; Tyrc; 1.

DR PROSITE; PS00022; EGF\_1; 3.  
 DR PROSITE; PS01186; EGF\_2; 3.  
 DR PROSITE; PS00026; EGF\_3; 1.  
 DR PROSITE; PS08053; FN3; 3.  
 DR PROSITE; PS08035; IG LIKE; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE\_TYR; 1.  
 KW ATP-binding; EGF-like domain; Glycoprotein; Immunoglobulin domain;  
 KW Kinase; Multigene family; Nucleotide-binding; Phosphorylation;  
 KW Receptor; Repeat; Signal; Transferrase; Transmembrane;  
 KW Tyrosine-protein kinase.  
 FT SIGNAL 1 18 Potential.  
 FT CHAIN 19 1122 Angiotensin 1 receptor.  
 FT TOPO\_DOM 19 744 Extracellular (Potential).  
 FT TRANSMEM 745 769 Potential.  
 FT TOPO\_DOM 770 1122 Cytoplasmic (Potential).  
 FT DOMAIN 44 123 Ig-like C2-type 1.  
 FT DOMAIN 210 252 EGF-like 1.  
 FT DOMAIN 254 299 EGF-like 2.  
 FT DOMAIN 301 341 EGF-like 3.  
 FT DOMAIN 350 440 Ig-like C2-type 2.  
 FT DOMAIN 445 535 Fibronectin type-III 1.  
 FT DOMAIN 542 632 Fibronectin type-III 2.  
 FT DOMAIN 638 730 Fibronectin type-III 3.  
 FT DOMAIN 822 1094 Protein kinase.  
 FT NP\_BIND 828 836 ATP (By similarity).  
 FT ACT\_SITE 962 962 Proton acceptor (By similarity).  
 FT BINDING 853 853 ATP (By similarity).  
 FT MOD\_RES 990 990 Phosphotyrosine (by autocatalysis) (By similarity).  
 FT CARBOHYD 140 140 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 158 158 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 399 399 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 438 438 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 464 464 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 558 558 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 595 595 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 648 648 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 690 690 N-linked (GlcNAc...) (Potential).  
 FT CONFLICT 161 171 PIHSPRHEVP -> LHPLSAPGMKYL (in Ref. 3).  
 FT CONFLICT 538 538 S -> C (in Ref. 1).  
 FT CONFLICT 736 736 A -> G (in Ref. 1 and 4).  
 FT CONFLICT 745 761 MLTILGSGACMTCITV -> DATHSHPWVWDFASPC (in Ref. 3).  
 FT CONFLICT 786 786 N -> NV (in Ref. 3 and 6).  
 FT CONFLICT 913 913 R -> G (in Ref. 3).  
 FT CONFLICT 925 931 AIANSTA -> CHRQOYS (in Ref. 3).  
 FT CONFLICT 1117 1117 S -> P (in Ref. 3).  
 SQ SEQUENCE 1122 AA; 125701 MW; F879623D103FFE96 CRC64;  
 Query Match 92.9%; Score 5584; DB 1; Length 1122;  
 Best Local Similarity 92.7%; Pred. No. 8.4e-312;  
 Matches 1043; Conservative 37; Mismatches 41; Indels 4; Gaps 3;  
 Qy 1 MDSIASLVLCVSLLSGTVEGAMDLLINSLPLVSDAETSLTCTIAGWRPHEPTIGRD 60  
 Db 1 MDSIAGLVLCVSLLSGTVEGAMDLLINSLPLVSDAETSLTCTIAGWHHEPTIGRD 60  
 Qy 61 FEALMNQHDPLEVTQDVTRWAKKVVYKREKASKINGAYFCEGRVGRGEAIRITMKMRQ 120  
 Db 61 FEALMNQHDPLEVTQDVTRWAKKVVYKREKASKINGAYFCEGRVGRGEAIRITMKMRQ 120  
 Qy 121 QASFLPATLTWVDKGDVNI SFKKVLKEBDVAYKNGSIHSPVPRHEVDILEVHLPH 180  
 Db 121 QASFLPATLTWVDKGDVNI SFKKVLKEBDVAYKNGSIHSPVPRHEVDILEVHLPH 180  
 Qy 181 APODAGVYSARYIGGNLFTSAFTRLVRRCEAKWGPCNHLCTACMNGVCHEDTCEC 240  
 Db 181 APODAGVYSARYIGGNLFTSAFTRLVRRCEAKWGPCNHLCTACMNGVCHEDTCEC 240  
 Qy 241 ICPFGFMGRTCEKACELHTFGTCKERCSGOGCKSVYFCLPDPYGCSCATWGKGLQNE 300





Db 479 VRPGEGBGHPVRFTTASIGLPPRGSLSLPSQALNLTWQPIFTSSSEDEYVEVE 538  
Qy 580 RRSVQ-KSDQNIKVPGNLTSLVLLNNLHPREQYVVRARVNTKQAGWSEDLTAWTL--SD 636  
Db 539 RWSQTRSDQNIKVPGNLTSLVLLNNLHPREQYVVRARVNTKQAGWSEELRAWTLDDSD 598  
Qy 637 ILPPQENIKINITHSSAVISWTILDGYSISSITIRYKVQGNKEDQVVDVKKNATTIQ 696  
Db 599 ILPPQENIKINITHSSAVISWTILDGYSISSITIRYKVQGNKEDQVVDVKKNATTIQ 658  
Qy 697 YOLKLEPETAYQVDIFAEENNIGSSNPAFSPHSLVLPESQAPADLGGGKMLLIALLGSAG 756  
Db 659 YOLKLEPETTVHVDIFAEENNIGSSNPAFSPHSLVLPESQAPADLGGGKMLLIALLGSAG 715  
Qy 757 MTCLTVLAFLLIOLKRAVORRMAQAFQNVREPAVQFNSGTLALNRKVQNPDPPIY 816  
Db 716 MTCTVLLAFLLIOLKRAVORRMAQAFQNVREPAVQFNSGTLALNRKVQNPDPPIY 775  
Qy 817 PVLWDNDIKFQDVGEGNFGVLRKIKKGLRMDAAIKRMKEYASKDHRDPFAGELEVL 876  
Db 776 PVLWDNDIKFQDVGEGNFGVLRKIKKGLRMDAAIKRMKEYASKDHRDPFAGELEVL 835  
Qy 877 CKLGHHPNIINLLGACEHGRGVLILAEYAPHGNLLDFLRKSRLVETDPAFAIANSTASTL 936  
Db 836 CKLGHHPNIINLLGACEHGRGVLILAEYAPHGNLLDFLRKSRLVETDPAFAIANSTASTL 895  
Qy 937 SSQQLLHFAADVARGMDVLSQKQFTHRDIAARNILVGENYVAKIADFGLSRGQEVYVKK 996  
Db 896 SSQQLLHFAADVARGMDVLSQKQFTHRDIAARNILVGENYVAKIADFGLSRGQEVYVKK 955  
Qy 997 MGRLPVRMAIESLNVSVYTTNSDWSYGVLLWEIVSLGGTPYCGMTCAELYEKLPQGYR 1056  
Db 956 MGRLPVRMAIESLNVSVYTTNSDWSYGVLLWEIVSLGGTPYCGMTCAELYEKLPQGYR 1015  
Qy 1057 LEKPLNCDDVYDLMRQWREKPYRSPFAQILVSLNRLMRLBKTYVNTTYLKEKTYAGI 1116  
Db 1016 LEKPLNCDDVYDLMRQWREKPYRSPFAQILVSLNRLMRLBKTYVNTTYLKEKTYAGI 1075  
Qy 1117 DCSAEAA 1124  
Db 1076 DCSAEAA 1083  
RESULT 10  
Qy Q80YS4\_MOUSE PRELIMINARY; PRT; 1072 AA.  
AC Q80YS4;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Tek protein.  
GN Name:Tek;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Embryo;  
EX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S.C., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Embryo;  
RA Strausberg R.;  
RL Submitted (APR-2003) to the EMBL/GenBank/DBSJ databases.  
DR EMBL; BC050824; AAH50824.1; -, mRNA.  
DR HSSP; Q02763; IFVR.  
DR SMR; Q80YS4; 761-1069.  
DR Ensembl; ENSMUSG00000006386; Mus musculus.  
DR MGI; MGI:98664; Tek.  
DR GO; GO:0005615; C:extracellular space; TAS.  
DR GO; GO:0016021; C:integral to membrane; TAS.  
DR GO; GO:0016337; P:cell-cell adhesion; IMP.  
DR GO; GO:007160; P:cell-matrix adhesion; IMP.  
DR GO; GO:0045765; P:regulation of angiogenesis; IMP.  
DR GO; GO:0030334; P:regulation of cell migration; IMP.  
DR InterPro; IPR000742; EGF\_2.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR003961; FN III.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR002049; Laminin EGF.  
DR InterPro; IPR000719; Prot kinase.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR InterPro; IPR008266; Tyr\_kinase\_AS.  
DR Pfam; PF00041; fn3; 3.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot kinase; 1.  
DR SMART; SM00180; EGF\_Lam; 1.  
DR SMART; SM00060; FN3\_3.  
DR SMART; SM00219; TyrKc; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_3.  
DR PROSITE; PS01186; EGF\_2; 3.  
DR PROSITE; PS00026; EGF\_3; 1.  
DR PROSITE; PS00853; FN3; 3.  
DR PROSITE; PS00835; IG LIKE; 1.  
DR PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN\_1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
SQ SEQUENCE 1072 AA; 120047 MW; E652FFB247D1ABB6 CRC64;  
Query Match 88.1%; Score 5296; DB 2; Length 1072;  
Best Local Similarity 88.3%; Pred. No. 2.7e-295;  
Matches 993; Conservative 37; Mismatches 41; Indels 54; Gaps 3;  
Qy 1 MDSLASLVLCGVSLLSGTVGEGAMDLLILINSLPLVSDAETSLTICIASGWRPHEPTIGRD 60  
Db 1 MDSLASLVLCGVSLLSGTVGEGAMDLLILINSLPLVSDAETSLTICIASGWRPHEPTIGRD 60  
Qy 61 FEALMNQHQDPLEVTQDVTRWAKKVVWVRKASKINGAYFCEGRVVRGEAIRITMKMRQ 120  
Db 61 FEALMNQHQDPLEVTQDVTRWAKKVVWVRKASKINGAYFCEGRVVRGEAIRITMKMRQ 120  
Qy 121 QASFLPATLTMTVDKGNVNIISFKKVLTKEDAVLYKNGSFTHSVPRHEVPDILEVHLPH 180  
Db 121 QASFLPATLTMTVDKGNVNIISFKKVLTKEDAVLYKNGSFTHSVPRHEVPDILEVHLPH 159  
Qy 181 AQPDAGVYSARYICGNLFTSAFTRLIVRRCEAQKNGPECNHLCTACMNGVYCHEDTGEC 240  
Db 160 -----CEAQKNGPDCSRPCTTCKNGVYCHEDTGEC 189  
Qy 241 ICPGFMGRTCEKACELHTFQRTCKERGSGGCKSYVFCFLPDPYGCSCATGWKGLQCNE 300  
Db 190 ICPGFMGRTCEKACELHTFQRTCKERGSGGCKSYVFCFLPDPYGCSCATGWKGLQCNE 249



```
QY 301 ACHPGFYGPDCKLRCSNNGENCDRFGQCLSPGMOGLQCEGREGIPRMTPTKIVDLPDHE 360
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
250 ACPGSGYGPDCKLRCHCTNEEICDRFGQCLSPGMOGLQCEGREGIPRMTPTKIVDLPDHE 309
QY 361 VNSGKFNPTCKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFVAFTTHRIILPPDSG 420
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
310 VNSGKFNPTCKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFVAFTTHRIILPPDSG 369
QY 421 VVCSVNTVAGVVEKPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPYFGDGPIKSKK 480
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
370 VVCSVNTVAGVVEKPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPYFGDGPIKSKK 429
QY 481 LLYKPVNHYEAWQHIQVNTNEIVTLNLEPRTEYELCVQLVRRGEHGGHGPVRRFTTAS 540
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
430 LLYKPVNHYEAWQHIQVNTNEIVTLNLEPRTEYELCVQLVRRGEHGGHGPVRRFTTAS 487
QY 541 IGLPPRGNLNPKSQTTNLNLTWQDIPFSSDDDFVEVERRSVQ-KSDQONIKVGNLTS 599
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
488 IGLPPRGNLNPKSQTTNLNLTWQDIPFSSDDDFVEVERRSVQ-KSDQONIKVGNLTS 547
QY 600 VLLNHLHPREQVYVVRARVNTKAQGEWSEDLTAWTILSDILPPOPENIKISNITHSSAVISW 659
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
548 VLLNHLHPREQVYVVRARVNTKAQGEWSEDLTAWTILSDILPPOPENIKISNITHSSAVISW 607
QY 660 TILDGYSISSITIRYKVGQNEQDQVVKIKNATIIQYOLKLEPETAYQVDIFAENNIG 719
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
608 TILDGYSISSITIRYKVGQNEQDQVVKIKNATIIQYOLKLEPETAYQVDIFAENNIG 667
QY 720 SSNPAPSHLVTLPESQAPADIGGCKMLLIALTGSAGMTCLTVLAFLLIQLKCANVQR 779
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
668 SSNPAPSHLVTLPESQAPADIGGCKMLLIALTGSAGMTCLTVLAFLLIQLKCANVQR 727
QY 780 RMAQAFQNVREPAVQFNSGTLALNRKVKQNPDPPIYVLDWNDIKFQDVIQEGNFGQVL 839
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
728 RMAQAFQNVREPAVQFNSGTLALNRKVKQNPDPPIYVLDWNDIKFQDVIQEGNFGQVL 787
QY 840 KARIKKGRLMDAAIKRMKEVASKDDHDFAGELEVLCKLGHHPNIINLLGACEHRGYLY 899
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
788 KARIKKGRLMDAAIKRMKEVASKDDHDFAGELEVLCKLGHHPNIINLLGACEHRGYLY 847
QY 900 LAIEVAPHGNIIDFLRKSRVLETPAPAIANSTASTLSSQQLLHPAADVARGMDVLSQK 959
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
848 LAIEVAPHGNIIDFLRKSRVLETPAPAIANSTASTLSSQQLLHPAADVARGMDVLSQK 907
QY 960 FTHRLAARNILVGENYVAKIADFGLSRGQEVYVVKTMGRPLVPRWMAIESLNYSVYTTNS 1019
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
908 FTHRLAARNILVGENYVAKIADFGLSRGQEVYVVKTMGRPLVPRWMAIESLNYSVYTTNS 967
QY 1020 DVWSYGVLLWEIVSLGGTPYCGMTCAELYKLPQGYRLEKPLNCDDEVDVLMRQCWRKBP 1079
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
968 DVWSYGVLLWEIVSLGGTPYCGMTCAELYKLPQGYRLEKPLNCDDEVDVLMRQCWRKBP 1027
QY 1080 YERPSPFAQLVSLNRLMRLBERKTYVNTLYEKTYAGIDCSABEA 1124
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1028 YERPSPFAQLVSLNRLMRLBERKTYVNTLYEKTYAGIDCSABEA 1072

RESULT 11
Q81XB8 HUMAN PRELIMINARY; PRT; 788 AA.
AC Q81XB8;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Receptor protein tyrosine kinase (Fragment).
GN Name-Tie-2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
```

```
RA Sata S., Nakagawa K., Yonemitsu Y., Onimaru M., Tani M., Teutsami N.,
RL Jin C., Sueishi K.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB086825; BAC45250.1; -, mRNA.
DR HSSP; P00740; 1EDM.
DR Ensembl; ENSG00000120156; Homo sapiens.
DR GO; GO:00046301; F:kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR00742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF00041; fn3; 3.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00060; FN3; 3.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00026; EGF_3; 1.
DR PROSITE; PS50853; FN3; 3.
DR PROSITE; PS50853; FN3; 3.
DR PROSITE; PS50835; IG Like; 1.
KW EGF-like domain; Immunoglobulin domain; Kinase; Receptor.
FT NON TER 788
SQ SEQUENCE 788 AA; 87622 MW; D33AA5F18408C237 CRC64;

Query Match 69.3%; Score 4168.5; DB 2; Length 788;
Best Local Similarity 94.8%; Pred. No. 9.7e-231;
Matches 788; Conservative 0; Mismatches 0; Indels 43; Gaps 1;

QY 1 MDLSASLVLCGVSLLSGTVEGAMDLLIINSLPLVSDAETSLTCTASGWRPHEPTIGRD 60
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MDLSASLVLCGVSLLSGTVEGAMDLLIINSLPLVSDAETSLTCTASGWRPHEPTIGRD 60

QY 61 FEALMNQHDPLEVTQDVTREWAKVWVKREKASKINGAYFCEGVRGAEIRITWKMQRQ 120
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 FEALMNQHDPLEVTQDVTREWAKVWVKREKASKINGAYFCEGVRGAEIRITWKMQRQ 120

QY 121 QASFLPATLTMTVDKGDVNIISFKVLIKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 QASFLPATLTMTVDKGDVNIISFKVLIKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180

QY 181 AQPQDAGYVSARYIGGNLFTSAFTRLIVRCEAQKMGPECHNLTCTACMNGVCHEDTGEC 240
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 AQPQDAGYVSARYIGGNLFTSAFTRLIVRCEAQKMGPECHNLTCTACMNGVCHEDTGEC 240

QY 241 ICPPGFMGRTCEKACELHTFGRTCKERCSGGEGCKSVYFCLPDPYGCSCATGKGLQCN 300
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 ICPPGFMGRTCEKACELHTFGRTCKERCSGGEGCKSVYFCLPDPYGCSCATGKGLQCN- 299

QY 301 ACHPGFYGPDCKLRCSNNGENCDRFGQCLSPGMOGLQCEGREGIPRMTPTKIVDLPDHE 360
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
300 -----EGIPRMTPTKIVDLPDHE 317

QY 361 VNSGKFNPTCKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFVAFTTHRIILPPDSG 420
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
318 VNSGKFNPTCKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFVAFTTHRIILPPDSG 377

QY 421 VVCSVNTVAGVVEKPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPYFGDGPIKSKK 480
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
378 VVCSVNTVAGVVEKPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPYFGDGPIKSKK 437

QY 481 LLYKPVNHYEAWQHIQVNTNEIVTLNLEPRTEYELCVQLVRRGEHGGHGPVRRFTTAS 540
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
438 LLYKPVNHYEAWQHIQVNTNEIVTLNLEPRTEYELCVQLVRRGEHGGHGPVRRFTTAS 497

QY 541 IGLPPRGNLNPKSQTTNLNLTWQDIPFSSDDDFVEVERRSVQSDQONIKVGNLTSV 600
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
498 IGLPPRGNLNPKSQTTNLNLTWQDIPFSSDDDFVEVERRSVQSDQONIKVGNLTSV 557

QY 601 LLLNHLHPREQVYVVRARVNTKAQGEWSEDLTAWTILSDILPPOPENIKISNITHSSAVISW 660
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
558 LLLNHLHPREQVYVVRARVNTKAQGEWSEDLTAWTILSDILPPOPENIKISNITHSSAVISW 617
```

Qy 661 ILDGYSISSITIRYKVOQKNEQHDVVKIKNATIIQYOLKLEPETAYQVDIFRAENNIGS 720  
 |||||  
 Db 618 ILDGYSISSITIRYKVOQKNEQHDVVKIKNATIIQYOLKLEPETAYQVDIFRAENNIGS 677  
 |||||  
 Qy 721 SNPAFSELVTLPSQAPADLGGGKMLLAIILSAGMTCLTVLLAFLIILQKRAVQRR 780  
 |||||  
 Db 678 SNPAFSELVTLPSQAPADLGGGKMLLAIILSAGMTCLTVLLAFLIILQKRAVQRR 737  
 |||||  
 Qy 781 MAQAFQNVREPAQFNSGTLALNRKVQKNPDPTIYPVLWDNDIKFQDVIG 831  
 |||||  
 Db 738 MAQAFQNVREPAQFNSGTLALNRKVQKNPDPTIYPVLWDNDIKFQDVIG 788  
 |||||

RESULT 12

TIE2_BRARE	STANDARD;	PRT; 1116 AA.
ID	TIE2_BRARE	
AC	O73791;	
DT	10-OCT-2003 (Rel. 42, Created)	
DT	10-OCT-2003 (Rel. 42, Last sequence update)	
DT	13-SEP-2005 (Rel. 48, Last annotation update)	
DE	Tyrosine-protein kinase receptor Tie-2 precursor (EC 2.7.1.112).	
GN	Name=tie2; Synonym=stie-2;	
OS	Brachydanio rerio (Zebrafish) (Danio rerio).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;	
OC	Cyprinidae; Danio.	
OX	NCBI_taxid=7955;	
RN	(1)	
RP	NUCLEOTIDE SEQUENCE, AND TISSUE SPECIFICITY.	
RC	TISSUE=Embryo;	
RX	MEDLINE=98264593; PubMed=9603430;	
RX	DOI=10.1002/(SICI)1097-0177(199805)212:1<133::AID-AJAL2>3.3.CO;2-3;	
RA	Lyons M.S., Bell B., Stainer D., Peters K.G.;	
RT	"Isolation of the zebrafish homologues for the tie-1 and tie-2 endothelium-specific receptor tyrosine kinases.";	
RL	Dev. Dyn. 212:133-140(1998).	
CC	-!- FUNCTION: Probably regulates endothelial cell proliferation, differentiation and guides the proper patterning of endothelial cells during blood vessel formation (By similarity).	
CC	-!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + a protein tyrosine phosphate.	
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein.	
CC	-!- TISSUE SPECIFICITY: Expressed in most populations of endothelial cells in 24 hours embryos. Not present in intersgmental vessels.	
CC	-!- SIMILARITY: Belongs to the Tyr protein kinase family. Tie subfamily.	
CC	-!- SIMILARITY: Contains 3 EGF-like domains.	
CC	-!- SIMILARITY: Contains 3 fibronectin type-III domains.	
CC	-!- SIMILARITY: Contains 2 Ig-like C2-type (immunoglobulin-like) domains.	

-----  
 This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.  
 -----  
 EMBL; AF053632; AAC09331.1; -; mRNA.  
 DR HSSP; Q02763; 1FVR.  
 DR SMR; O73791; 807-1113.  
 DR ZFIN; ZDB-GENE-990415-56; tie2.  
 DR InterPro; IPRO00742; EGF\_2.  
 DR InterPro; IPRO06209; EGF\_like.  
 DR InterPro; IPRO03961; FN\_III.  
 DR InterPro; IPRO06210; IEGF.  
 DR InterPro; IPRO03599; IG.  
 DR InterPro; IPRO07110; Ig-like.  
 DR InterPro; IPRO00719; Prot\_kinase.  
 DR InterPro; IPRO01245; Tyr\_kinase\_AS.  
 DR Pfam; PF00041; fn3; 3.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.

DR SMART; SM00181; EGF; 2.  
 DR SMART; SM00050; FN3; 3.  
 DR SMART; SM00409; IG; 1.  
 DR SMART; SM00219; TyrKc; 1.  
 DR PROSITE; PS00022; EGF\_1; 2.  
 DR PROSITE; PS01186; EGF\_2; 2.  
 DR PROSITE; PS00026; EGF\_3; 1.  
 DR PROSITE; PS00853; FN3; 3.  
 DR PROSITE; PS00835; IG\_LIKE; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
 KW ATP-binding; EGF-like domain; Glycoprotein; Immunoglobulin domain;  
 KW Kinase; Multigene family; Nucleotide-binding; Phosphorylation;  
 KW Receptor; Repeat; Signal; Transferase; Transmembrane;  
 KW Tyrosine-protein kinase.  
 FT SIGNAL 1 21 Potential.  
 FT CHAIN 22 1116 Tyrosine-protein kinase receptor Tie-2.  
 FT TOPO\_DOM 22 745 Extracellular (Potential).  
 FT TRANSMEM 746 766 Potential.  
 FT TOPO\_DOM 767 1116 Cytoplasmic (Potential).  
 FT DOMAIN 46 126 Ig-like C2-type 1.  
 FT DOMAIN 214 256 EGF-like 1.  
 FT DOMAIN 258 302 EGF-like 2.  
 FT DOMAIN 304 342 EGF-like 3.  
 FT DOMAIN 348 438 Ig-like C2-type 2.  
 FT DOMAIN 442 532 Fibronectin type-III 1.  
 FT DOMAIN 537 627 Fibronectin type-III 2.  
 FT DOMAIN 632 728 Fibronectin type-III 3.  
 FT NP\_BIND 822 830 Protein kinase.  
 FT ACT\_SITE 956 956 ATP (By similarity).  
 FT BINDING 847 847 ATP (By similarity).  
 FT MOD\_RES 984 984 Phosphotyrosine (by autocatalysis) (By similarity).  
 FT CARBOHYD 110 110 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 143 143 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 223 223 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 367 367 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 387 387 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 425 425 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 590 590 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 637 637 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 642 642 N-linked (GlcNAc... ) (Potential).  
 SQ SEQUENCE 1116 AA; 122361 MW; AA14EB9C745A8937 CRC64;  
 Query Match 48.5%; Score 2915.5; DB 1; Length 1116;  
 Best Local Similarity 51.6%; Pred. No. 1.3e-158;  
 Matches 596; Conservative 173; Mismatches 311; Indels 75; Gaps 25;

Qy 1 MDSLASIVLCVSLLLSGTVEGANDLILINSLPLVSD-AETSLTCIASGW----- 49  
 Db 4 LDSCATALLLG--CWMGSAVRISDVTLVNPDPVPSLPTAFSLICVSSDWSGGSVLALG 61  
 Qy 50 ---RPH-EPITIGRDFEALMNHQDPLEVTQDVTREWAKKVVWKKREKASKINGAYFCEG 104  
 Db 62 QEPFRPGSVLALQEP-----PTEPRPAAATVTWSSR-----SHAFGAFYQI 108  
 Qy 105 RVRGEAIRITMKVRQQAFLPATLTMTVDKGNVNISFKVLIKBEDAVIYKNGSFTHS 164  
 Db 109 R-NSGTRKIVTYKMLQEAFLPESLITITVNGENINISYSRLYSPEDTVIHKNGHFEHS 167  
 Qy 165 VPRHEVPDILEVHLPHAQPOD-AGVSNARYIGGNLFTSAFTRLIVRCEAKWGPCNHL 223  
 Db 168 SPKEDIISDIHYPTVNAKESHAIGIYAIRIISAAPSSAAITRLIVRSCRAFGWPGNCTES 227  
 Qy 224 CTACWNNGVCHETGECICPEGFMGRTECEKACELHFTGRTCKERCSCGEG-CKSVVFLCP 282  
 Db 228 CPCRANGVCDTGTGECVCPGFRGHTCDIVCGEGRFGAGCKERC--VDGVCRAIVFLCLR 285  
 Qy 283 DPYGCSCATGKGLQCNKAECHPGYPGDPCKLURCSNNGEMCDRFQGLCLSPGWQLQGER 342  
 Db 286 DPYGCSCASGWRGLSCNDACPDGYGAGCTQKVCYKAGR-CDRFRGCVCA-GRHGSRCBE 343



```

DR InterPro; IPR006209; EGF like.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00041; fn3; 3.
DR Pfam; PF00047; ig; 2.
DR PRINTS; PR0109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00060; FN3; 3.
DR SMART; SM00409; IG; 1.
DR SMART; SM00219; Tyrg; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS00026; EGF_3; 3.
DR PROSITE; PS00853; FN3; 3.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Direct protein sequencing; EGF-like domain; Glycoprotein;
KW Immunoglobulin domain; Kinase; Multigene family; Nucleotide-binding;
KW Phosphorylation; Receptor; Repeat; Signal; Transmembrane;
KW Tyrosine-protein kinase.
FT SIGNAL 1 21
FT CHAIN 22 1138
FT TOPO_DOM 22 759
FT TRANSMEM 760 784
FT TOPO_DOM 785 1138
FT DOMAIN 43 105
FT DOMAIN 214 256
FT DOMAIN 258 303
FT DOMAIN 305 345
FT DOMAIN 372 426
FT DOMAIN 446 537
FT DOMAIN 545 637
FT DOMAIN 644 736
FT DOMAIN 839 1118
FT NP_BIND 845 853
FT ACT_SITE 979 979
FT BINDING 870 870
FT MOD_RES 1007 1007
FT CARBOHYD 83 83
FT CARBOHYD 161 161
FT CARBOHYD 503 503
FT CARBOHYD 596 596
FT CARBOHYD 709 709
SQ SEQUENCE 1138 AA; 125090 MW; 3B42BE33678C58A1 CRC64;

Query Match
Best Local Similarity 42.0%; Score 2527; DB 1; Length 1138;
Matches 541; Conservative 165; Mismatches 372; Indels 80; Gaps 27;

QY 14 LLLSGTVEGAMDILILNSLPLVSDAETSLTCT-----ASGWRPHEPTITIGRDFEAL 64
DB 13 LFLASHGAAVDLTLLANRLTDPQRFELTCTVSGEAGAGRGSDAWGP--PLLLEKDDRV 70

QY 65 MNQODPLEVTDVTRWAKVWVR--EKASKINGAYCEGEVRGEATRIRTMKRQOAS 123
DB 71 RTPPGPLR-----LARGNSHQVTLRGFKPSDLVGVFCVGAGARRTRVTVHNSPGA 126

QY 124 FLPATLTMTVDKDNVNTSFKKVLTKEDAVTYKNGSFTHSVPRHEVPD---ILEVHLPH 180
DB 127 LLPDKVTHVNGDITAVLSARVHKQKQTDVIWKSNGSYFYTLDWHEAQDGRFLQ--LPN 184

QY 181 AQPDAGVYSARYIGNLTFTSFTRLIVRRCBAQKRGFECHNLCTACMNGVGVCHDTEGEC 240
DB 185 VQPPSSGIYSATYLEASPLGSAFFRLIVRGCGAGRWGPGCTKECPGLHGGVCHDHGEC 244

```

```

QY 241 ICPGPFMRTCEKACELHTFORTCKERCSGOEGCKSYVFCPLPDPYGCSCATGWKLCQNE 300
DB 245 VCPFGFTGTGRCEQACREGRFGQSCQEQCPGTSRGGLTFCFLPDPYGCSCGSGWRGSCQE 304

QY 301 ACHPGFYGPDCKLRCSNNGMCDRFOGLCSPGQGLQCEREGIPRMTPIVDLPDIE 360
DB 305 ACARHFGADCLQCCQNGTCDRFSGCVCPGSHGVHCEKSD---RIPQILNWSLE 361

QY 361 VNSGKFNDI-CKASGWLPTNEMTLVKPDGTVLHPKDFNHTDHFSAVIFTHRLPPDS 419
DB 362 FNLETMPRINCAAGNPPVRGSIELARKPDCTVLLSTKAI VEPEKTTAEFVPRVLVADS 421

QY 420 GWNVCSVNTVAGMVKPENSIVKVLPRPLANPNVIDTCHNEAVIN--ISSEPYGDDGPIK 477
DB 422 GFWECRVSTSGQDSRRFKVNVKVPVPPVLAAPRLTLTKQSRQLVVSPLVS---FSGDGPIS 478

QY 478 SKLLYKPVNHYEAWQHIQV--TNEIVTINYLEPRTEYELCVQLVRRGEGGSGHPQVRRF 536
DB 479 TVRLHYRPQDSTMDWSTIVDPSENVTLMNLRPKTYGSVRVQLSRPGCGEGGAGPPTLM 538

QY 537 TTASIGLPPPRGLNLLP-----KSQTTMLTWQ-PIFPSS--EDDPYVEVERRSVOKS 586
DB 539 TT---DCPEPL---LQPWLEGWHEGTDRLAVSNLSPLVPGPLVGDGFLRLMDGTRQE 592

QY 587 DQONIKVPGNLTSLVLLNHLHPREOVVVRARV-NTKACQGEWSEDLTAWTSLDILPPQENI 645
DB 593 RRENVSPQARTA--LLTGLTPGTHYQLDVQVYHCTLLGPASPAPAHV--LLPSPGPPAPRHL 650

QY 646 KISNITHSSAIVSW---TILDGYSISSITIRYKQGNEDQ---HVDVKIKNATIIQYQL 699
DB 651 HAQALSDEIQLTWKHPEALPG-PISKYVVEVQVAGGAGDPLWDVDVDRPEETSII---- 705

QY 700 KGLSPETAYQVDIFA-----ENNIGSSNPAPSHLVTLPESQAPADLGGQKM 746
DB 706 RGLNASTRYLFRMRASIQGLGDWNSNTVEESTLGNGLQAEQ---PVQBSRA-ABEGLDQ 760

QY 747 LLAILGSGMTCLTVLLAFLIILQLKRVORVMAQAFQV-VREEPVAVQNSGTLALNR 805
DB 761 LILAVGSGVSATCLTILAAALTLCIRRSCLHRRRTFTYQSGSGEETTLQSSGTLTTR 820

QY 806 KVXGNPDPTIYPVLWDNDIKFDQVIGEGNFQVILKARIKKDGLRMDAAIKRMKEYASKDD 865
DB 821 RPKLQPELSYPVLEWEDITFEDLIGEGNFQVIRAMIKKDGKMNAAIKMLKEYASEND 880

QY 866 HRDFAGELEVLCGLHHPNIIINLIGACHERGYLYLAIEYAPHGNLLDFLRKSRVLETPA 925
DB 881 HRDFAGELEVLCGLHHPNIIINLIGACKNKGYLIAIEYAPYGNLLDFLRKSRVLETPA 940

QY 926 FAIRNSTASTLSSQOLLHFAADVARGMDYLSQOFIHRDLAARNILVGENYVAKIADFGL 985
DB 941 FAREHGTASTLSSQOLLRFASDAANGQYLSQOFIHRDLAARNVVLGENLASKIADFGL 1000

QY 986 SRGOEVYKKTMTGRPLVRWMAIESLNTSVYTTNSDVMSYGYLLWEIVSLGTPYCGMTCA 1045
DB 1001 SRGEVYVKTMTGRPLVRWMAIESLNTSVYTTKSDVMSFGVLLWEIVSLGTPYCGMTCA 1060

QY 1046 ELYEKLPGCYRLEKPLNCDDEVYDLMRQCWEKPERPSPAQILVSLNRMLEERTYNT 1105
DB 1061 ELYEKLPGCYRMEQPRNCDDEVYELMRQCWRDRPYERPPPAQIALQLGRMLEARKAYNM 1120

QY 1106 TLYEKFTYAGIDCSAEEA 1123
DB 1121 SLFENFTYAGIDATBEEA 1138

```

```

RESULT 14
TIE1_BOVIN
ID TIE1_BOVIN STANDARD; PRT; 1136 AA.
AC Q06805;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)

```



Db 694 DRPEETSTIV----RGLNASTRYLFRVRASVQPGDWSNVVEQSTLGNGLQIEGSPVQBIIH 749  
QY 738 PADLGGGKMLIAIIGSAGMTCLTVLLAFLIILQLKRANVOREMAQAFQN-VREBPVQVF 796  
Db 750 AABEGLDQOLVLAVGVGSATCLTILAALLTLACIRKSLCHRRRTFTYOSGSEETILQF 809  
QY 797 NSGTLLALNRKVNPNPDPTIYPVLDWNDIKFDQVIGEGNFGQVLKARIKKDGLRMDAAIKR 856  
Db 810 SSGTLTLTRPKPQPELPYNPVLWEEDITFEDLIGEGNFGQVIRAWIKKGLKWNAAIKM 869  
QY 857 MKEVASKDDHRDFAGELEVLCKLGHHPNIINLLGACHERGVLYLAIYVAPGNLLDFLRK 916  
Db 870 LKEYASNDHRDFAGELEVLCKLGHHPNIINLLGACENRGVLYLAIYVAPYCNLLDFLRK 929  
QY 917 SRVLTFDPAFAIANSTASTLSSQQLLHFAADVARGMDVLSQKQFTHRDLAARNLIVAGNY 976  
Db 930 SRVLTFDPAFAREHGTASTLSRQLRFPASDAANGQVLSKQFTHRDLAARNLIVAGNL 989  
QY 977 VAKIADFGLSRGQEVYVKTWGRLPVRWMAIESLNSVYTTNSDVWSYGVLLWEIVSLGG 1036  
Db 990 ASKIADFGLSRGQEVYVKTWGRLPVRWMAIESLNSVYTTKSDVWSYGVLLWEIVSLGG 1049  
QY 1037 TPYCGMTCAELYEKLPOQYRLKLEKPLNCDDDEVYDLMRQCWRKPYRPSFAQILVSLNRM 1096  
Db 1050 TPYCGMTCAELYEKLPOQYRMEQPRNCDDDEVYELMRQCWRDPRYRPPFAQIALQLGRML 1109  
QY 1097 EERTYVNTTLYEKTYPYAGIDCSAEEA 1123  
Db 1110 EARKAYVNNLSUFENPTYAGIDATAEEA 1136  
  
RESULT 15  
ID Q8BG12\_MOUSE PRELIMINARY; PRT: 1134 AA.  
AC Q8BG12;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DE Mus musculus 13 days embryo heart cDNA, RIKEN full-length enriched  
DE library, clone:D330005P09 product:tyrosine kinase receptor 1, full  
DE insert sequence (Mus musculus 13 days embryo lung cDNA, RIKEN full-  
DE length enriched library, clone:D43008P04 product:tyrosine kinase  
DE receptor 1, full insert sequence).  
GN Name=Tiel;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Heart, and Lung;  
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning.";  
RL Meth. Enzymol. 303:19-44(1999).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Heart, and Lung;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo S., Nikaide I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kotsuki S.,  
RT Hayashizaki Y.;  
RL "Functional annotation of a full-length mouse cDNA collection.";  
RN Nature 409:685-690(2001).  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Heart, and Lung;  
RX MEDLINE=20354683; PubMed=12466851; DOI=10.1038/nature01266;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
RA Kanaai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
RA Maglott D.R., Maltas L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,  
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs";  
RL Nature 420:563-573(2002).  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Heart, and Lung;  
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
RN [5]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Heart, and Lung;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagao K., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
RN [6]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Heart, and Lung;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
RA Hayashida K., Hayatsu N., Hitamoto K., Hiraoka T., Hirozane T.,  
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,







GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2006, 01:48:13 ; Search time 50 Seconds  
(without alignments)  
1858.549 Million cell updates/sec

Title: US-10-633-742-8  
Perfect score: 6013  
Sequence: 1 MDLSASLVLCGVSLLSGTV.....TTLXKFTYAGIDCSABAA 1124

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/prodata/1/aaa/5 COMB.pep.\*
- 2: /cgn2\_6/prodata/1/aaa/6 COMB.pep.\*
- 3: /cgn2\_6/prodata/1/aaa/H COMB.pep.\*
- 4: /cgn2\_6/prodata/1/aaa/PCTUS COMB.pep.\*
- 5: /cgn2\_6/prodata/1/aaa/RE COMB.pep.\*
- 6: /cgn2\_6/prodata/1/aaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6013	100.0	1124	1 US-08-323-474-2	Sequence 2, Appli
2	6013	100.0	1124	2 US-09-949-016-5946	Sequence 5946, Ap
3	6013	100.0	1124	4 PCT-US993-06093-2	Sequence 2, Appli
4	6008	99.9	1157	2 US-09-949-016-9568	Sequence 9568, Ap
5	5575	92.7	1122	1 US-08-278-089A-6	Sequence 6, Appli
6	5575	92.7	1122	1 US-08-838-957A-6	Sequence 6, Appli
7	5542	92.2	1118	1 US-07-934-393B-2	Sequence 2, Appli
8	5542	92.2	1118	1 US-08-278-089A-2	Sequence 2, Appli
9	5542	92.2	1118	1 US-08-838-957A-2	Sequence 96, Appli
10	5518	91.8	1101	1 US-08-469-537A-96	Sequence 1, Appli
11	4043	67.2	977	2 US-09-590-656-1	Sequence 1, Appli
12	4043	67.2	977	2 US-09-733-764-1	Sequence 2, Appli
13	2623	43.6	704	2 US-09-590-656-2	Sequence 2, Appli
14	2623	43.6	704	2 US-09-733-764-2	Sequence 2, Appli
15	2527	42.0	1138	1 US-08-220-240A-5	Sequence 5, Appli
16	2527	42.0	1150	2 US-09-949-016-7769	Sequence 7769, Ap
17	2511	41.8	1138	1 US-08-323-474-8	Sequence 8, Appli
18	2511	41.8	1138	1 US-08-469-537A-98	Sequence 98, Appli
19	2499.5	41.6	1135	1 US-08-469-537A-97	Sequence 97, Appli
20	1582	26.3	301	1 US-07-934-393B-4	Sequence 4, Appli
21	1582	26.3	301	1 US-08-278-089A-4	Sequence 4, Appli
22	1582	26.3	301	1 US-08-278-089A-14	Sequence 14, Appli
23	1582	26.3	301	1 US-08-469-537A-99	Sequence 99, Appli
24	1582	26.3	301	1 US-08-838-957A-4	Sequence 4, Appli
25	1309	21.8	291	1 US-08-701-191A-22	Sequence 22, Appli
26	1309	21.8	291	2 US-09-664-526-22	Sequence 22, Appli
27	1235	20.5	239	1 US-08-323-474-3	Sequence 3, Appli

28	1235	20.5	239	4 PCT-US993-06093-3	Sequence 3, Appli
29	701.5	11.7	820	1 US-08-166-717D-6	Sequence 6, Appli
30	695.5	11.6	820	1 US-07-921-807B-3	Sequence 3, Appli
31	695.5	11.6	820	1 US-08-441-944A-3	Sequence 3, Appli
32	695.5	11.6	820	2 US-08-439-992A-1	Sequence 1, Appli
33	695.5	11.6	820	2 US-09-620-561-1	Sequence 1, Appli
34	694.5	11.5	822	1 US-07-997-133-1	Sequence 1, Appli
35	694.5	11.5	822	1 US-07-921-807B-4	Sequence 4, Appli
36	694.5	11.5	822	1 US-08-441-944A-4	Sequence 4, Appli
37	694.5	11.5	822	1 US-08-451-822A-12	Sequence 12, Appli
38	694.5	11.5	822	2 US-08-439-992A-2	Sequence 2, Appli
39	694.5	11.5	822	2 US-08-323-430-12	Sequence 12, Appli
40	694.5	11.5	822	2 US-09-620-561-2	Sequence 2, Appli
41	693	11.5	795	2 US-09-949-016-7119	Sequence 7119, Ap
42	691	11.5	806	2 US-09-383-630-3	Sequence 3, Appli
43	687.5	11.4	822	1 US-08-459-296-2	Sequence 2, Appli
44	683.5	11.4	891	2 US-09-345-473E-25	Sequence 25, Appli
45	683.5	11.4	891	2 US-09-862-027-25	Sequence 25, Appli

ALIGNMENTS

RESULT 1  
US-08-323-474-2  
; Sequence 2, Application US/08323474  
; Patent No. 5447860  
; GENERAL INFORMATION:  
; APPLICANT: Ziegler, Steven F.  
; TITLE OF INVENTION: NOVEL TYROSINE KINASE  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/323,474  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/905,600  
; FILING DATE: 26-JUN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Seese, Kathryn A.  
; REGISTRATION NUMBER: 32,172  
; REFERENCE/DOCKET NUMBER: 2609  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 233-0644  
; TELEX: 756822  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1124 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-323-474-2

Query Match 100.0%; Score 6013; DB 1; Length 1124;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDLSASLVLCGVSLLSGTVGAMDLILINSPLVSDAETSLTCTASGWRPHEPTIGRD 60  
DB 1 MDLSASLVLCGVSLLSGTVGAMDLILINSPLVSDAETSLTCTASGWRPHEPTIGRD 60

```
QY 61 FEALMNQHDPLEVTDVTRWAKKVVWREKASKINGAYFCEGRVGEAIBIRTMKMRQ 120
Db 61 FEALMNQHDPLEVTDVTRWAKKVVWREKASKINGAYFCEGRVGEAIBIRTMKMRQ 120
QY 121 QASFLPATLTMTVDKGDVNIISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180
Db 121 QASFLPATLTMTVDKGDVNIISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180
QY 181 AQPQAGVVSARYIGGNLFTSAFTRLIVRRCEAQKWGPCNHLCTACMNGVCHEDTGEC 240
Db 181 AQPQAGVVSARYIGGNLFTSAFTRLIVRRCEAQKWGPCNHLCTACMNGVCHEDTGEC 240
QY 241 ICPPGFMGRTCEKACELHTFGTCKERCSCGGCKSYVFCCLDPYGCSCATGWKGLQCN 300
Db 241 ICPPGFMGRTCEKACELHTFGTCKERCSCGGCKSYVFCCLDPYGCSCATGWKGLQCN 300
QY 301 ACHPGFYGPDCKLRCSNNGEMCDRFGQCLCSPGWQGLQCEREGIPRMTPKIVDLPDHTE 360
Db 301 ACHPGFYGPDCKLRCSNNGEMCDRFGQCLCSPGWQGLQCEREGIPRMTPKIVDLPDHTE 360
QY 361 VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAIFTIHRILPPDSG 420
Db 361 VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAIFTIHRILPPDSG 420
QY 421 VVWCSVNTVAGVVEKPFNISVKVLPKLNAPNVIDTGHNFVAVINISSEPYFGDGP1KSKK 480
Db 421 VVWCSVNTVAGVVEKPFNISVKVLPKLNAPNVIDTGHNFVAVINISSEPYFGDGP1KSKK 480
QY 481 LLYKPNVHYEAWQHIOVTNEIVTLNLEPRTEYELCVQLVRREGEGEGHPVRRFTTAS 540
Db 481 LLYKPNVHYEAWQHIOVTNEIVTLNLEPRTEYELCVQLVRREGEGEGHPVRRFTTAS 540
QY 541 IGLPPRGNLNLPKSGOTTLNLTWQPIFPSSDDFYVEVERRSVQKSDQONIKVPGNLTSV 600
Db 541 IGLPPRGNLNLPKSGOTTLNLTWQPIFPSSDDFYVEVERRSVQKSDQONIKVPGNLTSV 600
QY 601 LNNLHPREQYVVRVAVNTKAQGESEDLTAWTSLDILPQPENIKISNI THSSAVISWT 660
Db 601 LNNLHPREQYVVRVAVNTKAQGESEDLTAWTSLDILPQPENIKISNI THSSAVISWT 660
QY 661 ILDGYSISSITIRYKQGNEDQVDVKIKNATIIQYQLKLEPETAYQVDFI AENNIGS 720
Db 661 ILDGYSISSITIRYKQGNEDQVDVKIKNATIIQYQLKLEPETAYQVDFI AENNIGS 720
QY 721 SNPAFSHELVTLPESQAPADLGGGKMLLIIAILGSGMTCLTVLLAPLIILQIKRANVQR 780
Db 721 SNPAFSHELVTLPESQAPADLGGGKMLLIIAILGSGMTCLTVLLAPLIILQIKRANVQR 780
QY 781 MAQAFONVREEPVAVQNSGTLALNRKVKNPDPPTIYPVLDMNDIKFQDVI GRNGFQVLK 840
Db 781 MAQAFONVREEPVAVQNSGTLALNRKVKNPDPPTIYPVLDMNDIKFQDVI GRNGFQVLK 840
QY 841 ARIKKDGLRMDAAIKRMKEYASKODHRDPAGELEVLCKLGHHPNIIINLIGACEHRGYLYL 900
Db 841 ARIKKDGLRMDAAIKRMKEYASKODHRDPAGELEVLCKLGHHPNIIINLIGACEHRGYLYL 900
QY 901 AIEYAPHGNLLDFLRKSRVLETDPAFAIANSTASTLSSQQLLHFAADVARGMDYLSQKOF 960
Db 901 AIEYAPHGNLLDFLRKSRVLETDPAFAIANSTASTLSSQQLLHFAADVARGMDYLSQKOF 960
QY 961 IHRDLAARNILYGENYVAKIADPGLSRGOVYVYKTMGRLPVRWMAIESLINSVYTTNSD 1020
Db 961 IHRDLAARNILYGENYVAKIADPGLSRGOVYVYKTMGRLPVRWMAIESLINSVYTTNSD 1020
QY 1021 VNSYGVLLWEIYSLGTPYCGMTCAELYEKLPGYRLEKPLNCDDEYDLMRCQWREKPY 1080
Db 1021 VNSYGVLLWEIYSLGTPYCGMTCAELYEKLPGYRLEKPLNCDDEYDLMRCQWREKPY 1080
QY 1081 ERPSFAQIILVSNLRMLEERKTYNTTLYEKFTYVAGIDCSAEAA 1124
Db 1081 ERPSFAQIILVSNLRMLEERKTYNTTLYEKFTYVAGIDCSAEAA 1124
```

```
RESULT 2
US-09-949-016-5946
; Sequence 5946, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5946
; LENGTH: 1124
; TYPE: PR1
; ORGANISM: Human
US-09-949-016-5946
```

```
Query Match 100.0%; Score 6013; DB 2; Length 1124;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MDSLASLVLCGVSLLSGTVEGAMDILLINSIPLVSDAETSLTCIASGWRPHEPITIGRD 60
Db 1 MDSLASLVLCGVSLLSGTVEGAMDILLINSIPLVSDAETSLTCIASGWRPHEPITIGRD 60
QY 61 FEALMNQHDPLEVTDVTRWAKKVVWREKASKINGAYFCEGRVGEAIBIRTMKMRQ 120
Db 61 FEALMNQHDPLEVTDVTRWAKKVVWREKASKINGAYFCEGRVGEAIBIRTMKMRQ 120
QY 121 QASFLPATLTMTVDKGDVNIISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180
Db 121 QASFLPATLTMTVDKGDVNIISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180
QY 181 AQPQAGVVSARYIGGNLFTSAFTRLIVRRCEAQKWGPCNHLCTACMNGVCHEDTGEC 240
Db 181 AQPQAGVVSARYIGGNLFTSAFTRLIVRRCEAQKWGPCNHLCTACMNGVCHEDTGEC 240
QY 241 ICPPGFMGRTCEKACELHTFGTCKERCSCGGCKSYVFCCLDPYGCSCATGWKGLQCN 300
Db 241 ICPPGFMGRTCEKACELHTFGTCKERCSCGGCKSYVFCCLDPYGCSCATGWKGLQCN 300
QY 301 ACHPGFYGPDCKLRCSNNGEMCDRFGQCLCSPGWQGLQCEREGIPRMTPKIVDLPDHTE 360
Db 301 ACHPGFYGPDCKLRCSNNGEMCDRFGQCLCSPGWQGLQCEREGIPRMTPKIVDLPDHTE 360
QY 361 VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAIFTIHRILPPDSG 420
Db 361 VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAIFTIHRILPPDSG 420
QY 421 VVWCSVNTVAGVVEKPFNISVKVLPKLNAPNVIDTGHNFVAVINISSEPYFGDGP1KSKK 480
Db 421 VVWCSVNTVAGVVEKPFNISVKVLPKLNAPNVIDTGHNFVAVINISSEPYFGDGP1KSKK 480
QY 481 LLYKPNVHYEAWQHIOVTNEIVTLNLEPRTEYELCVQLVRREGEGEGHPVRRFTTAS 540
Db 481 LLYKPNVHYEAWQHIOVTNEIVTLNLEPRTEYELCVQLVRREGEGEGHPVRRFTTAS 540
QY 541 IGLPPRGNLNLPKSGOTTLNLTWQPIFPSSDDFYVEVERRSVQKSDQONIKVPGNLTSV 600
Db 541 IGLPPRGNLNLPKSGOTTLNLTWQPIFPSSDDFYVEVERRSVQKSDQONIKVPGNLTSV 600
QY 601 LNNLHPREQYVVRVAVNTKAQGESEDLTAWTSLDILPQPENIKISNI THSSAVISWT 660
Db 601 LNNLHPREQYVVRVAVNTKAQGESEDLTAWTSLDILPQPENIKISNI THSSAVISWT 660
```



```
Db 961 IHRDLAARNILVGENYVAKIADFGLSRGQEVVYKKTGRLPVRWMAIESLNSVYVTTNSD 1020
Qy 1021 VMSYGVLLWEIVSLGTPYCGMTCAELYEKLPQGYRLEKPLNCDDEVYDLMRQCWREKPY 1080
Db 1021 VMSYGVLLWEIVSLGTPYCGMTCAELYEKLPQGYRLEKPLNCDDEVYDLMRQCWREKPY 1080
Qy 1081 ERPSFAQILVSLNRLMEERKTYVNTLYEKFTYAGIDCSABEAA 1124
Db 1081 ERPSFAQILVSLNRLMEERKTYVNTLYEKFTYAGIDCSABEAA 1124

RESULT 4
US-09-949-016-9568
; Sequence 9568, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9568
; LENGTH: 1157
; TYPE: PR
; ORGANISM: Human
US-09-949-016-9568

Query Match 99.9%; Score 6008; DB 2; Length 1157;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1123; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MDSLASLVLCGVSLLSGTVEGAMDILILNSPLVSDAETSITCISAGWRPHEPTIGRD 60
Db 34 MDSLASLVLCGVSLLSGTVEGAMDILILNSPLVSDAETSITCISAGWRPHEPTIGRD 93

Qy 61 PEALMNQODPLEVTDQVTRWAKVWVKREKASKINGAYFCEGVRGFEAIRTMKWRQ 120
Db 94 PEALMNQODPLEVTDQVTRWAKVWVKREKASKINGAYFCEGVRGFEAIRTMKWRQ 153

Qy 121 QASFLPATLTMVDKGDVNIISFKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180
Db 154 QASFLPATLTMVDKGDVNIISFKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 213

Qy 181 AQPQDAGVYSARYIGNLFTSAFTRLIVRRCBAQKWGPCNHLCTACMNNGVCHEDTGEC 240
Db 214 AQPQDAGVYSARYIGNLFTSAFTRLIVRRCBAQKWGPCNHLCTACMNNGVCHEDTGEC 273

Qy 241 ICPPFGMGTCEKACELHTFGTTCERCSGGQCKSVYFCLPDPYGCSCATGWKGLQNE 300
Db 274 ICPPFGMGTCEKACELHTFGTTCERCSGGQCKSVYFCLPDPYGCSCATGWKGLQNE 333

Qy 301 ACHPGFYGPDCKLRKSCNNGECMDRFQGLCSFGWQGLQCEGIPRMTPKIVDLPDHIE 360
Db 334 ACHPGFYGPDCKLRKSCNNGECMDRFQGLCSFGWQGLQCEGIPRMTPKIVDLPDHIE 393

Qy 361 VNSGKFNDICKASGWLPLTNBEMTIVKPDGTVLHPKDFNHTDHFSAIFTTHRIILPPDSG 420
Db 394 VNSGKFNDICKASGWLPLTNBEMTIVKPDGTVLHPKDFNHTDHFSAIFTTHRIILPPDSG 453

Qy 421 VMVCSVNTVAGWVEKPNISVKVLPKPNAPNVIDTGNFAVINISSESPYFGDGPISKK 480
Db 454 VMVCSVNTVAGWVEKPNISVKVLPKPNAPNVIDTGNFAVINISSESPYFGDGPISKK 513

Qy 481 LLYKPVNHYEAQH1QVNTNEIVTLNLEPRTEYELCVQLVRRGEGEGHGPVRRFTTAS 540
```

```
Db 514 LLYKPVNHYEAQH1QVNTNEIVTLNLEPRTEYELCVQLVRRGEGEGHGPVRRFTTAS 573
Qy 541 IGLPPRGINLLPKSQTTILNTWQDIPFSSSEDDDFYVEVERRSVQKSDOONIKVPGNLTSV 600
Db 574 IGLPPRGINLLPKSQTTILNTWQDIPFSSSEDDDFYVEVERRSVQKSDOONIKVPGNLTSV 633
Qy 601 LLNLHPREQYVVRARVNTKAGQWSEDLTAWTLDILPPOPENIKISINTHSSAVISWT 660
Db 634 LLNLHPREQYVVRARVNTKAGQWSEDLTAWTLDILPPOPENIKISINTHSSAVISWT 693
Qy 661 ILDGYSSISITIRYKVOGKNEHQHVDVKIKNATIIQYQIKGLEPETAYQVDFIAENNIGS 720
Db 694 ILDGYSSISITIRYKVOGKNEHQHVDVKIKNATIIQYQIKGLEPETAYQVDFIAENNIGS 753
Qy 721 SNPAFSELVTLPEAQADLGGGKMLLIIILGSAGMTCLTVLLAFLIILQKRVQR 780
Db 754 SNPAFSELVTLPEAQADLGGGKMLLIIILGSAGMTCLTVLLAFLIILQKRVQR 813
Qy 781 MAQAFQNVREBPVAVQFNSGTILALNRKVNNDPTIYPVLDMNDIKFQDVIGRGNFGVLIK 840
Db 814 MAQAFQNVREBPVAVQFNSGTILALNRKVNNDPTIYPVLDMNDIKFQDVIGRGNFGVLIK 873
Qy 841 ARIKKGGLRMDAAIKRMKEYASKODHRDFAGELEVLCKLGHHPNIINLLGACEHGYLYL 900
Db 874 ARIKKGGLRMDAAIKRMKEYASKODHRDFAGELEVLCKLGHHPNIINLLGACEHGYLYL 933
Qy 901 AIEYAPHGNLLDFLRKSRVLETDPAFAJANSTASTLSSQQLLHFAADVARGMDYLSQKQF 960
Db 934 AIEYAPHGNLLDFLRKSRVLETDPAFAJANSTASTLSSQQLLHFAADVARGMDYLSQKQF 993
Qy 961 IHRDLAARNILVGENYVAKIADFGLSRGQEVVYKKTGRLPVRWMAIESLNSVYVTTNSD 1020
Db 994 IHRDLAARNILVGENYVAKIADFGLSRGQEVVYKKTGRLPVRWMAIESLNSVYVTTNSD 1053
Qy 1021 VMSYGVLLWEIVSLGTPYCGMTCAELYEKLPQGYRLEKPLNCDDEVYDLMRQCWREKPY 1080
Db 1054 VMSYGVLLWEIVSLGTPYCGMTCAELYEKLPQGYRLEKPLNCDDEVYDLMRQCWREKPY 1113
Qy 1081 ERPSFAQILVSLNRLMEERKTYVNTLYEKFTYAGIDCSABEAA 1124
Db 1114 ERPSFAQILVSLNRLMEERKTYVNTLYEKFTYAGIDCSABEAA 1157

RESULT 5
US-08-278-089A-6
; Sequence 6, Application US/08278089A
; Patent No. 5681714
; GENERAL INFORMATION:
; APPLICANT: Breitman, Martin L.
; APPLICANT: Roseant, Janet
; APPLICANT: Dumont, Daniel J.
; APPLICANT: Yamaguchi, Terry P.
; TITLE OF INVENTION: No. 5681714el Receptor Tyrosine Kinase
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/278,089A
; FILING DATE: 20-JUL-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurdydk, Linda M.
```

```
;
; REGISTRATION NUMBER: 34,971
; REFERENCE/DOCKET NUMBER: 3153-111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1122 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-278-089A-6

Query Match 92.7%; Score 5575; DB 1; Length 1122;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1041; Conservative 37; Mismatches 43; Indels 4; Gaps 3;

QY 1 MDSLASLVLCGVSLLLSGTVEGMDLILINSPLVSDAETSLTCTASGWRPHEPTIGRD 60
DB 1 MDSLAGLVLCGVSLLLYGVVEGMDLILINSPLVSDAETSLTCTASGWHPEPTIGRD 60

QY 61 FEALNNQHDPLEVTQDVTREWAKKVVWKKREKASKINGAYFCEGRVGRGEAIRITMKMRQ 120
DB 61 FEALNNQHDPLEVTQDVTREWAKKVVWKKREKASKINGAYFCEGRVGRGEAIRITMKMRQ 120

QY 121 QASFLPATLTMTVDKGDNNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180
DB 121 QASFLPATLTMTVDKGDNNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180

QY 181 APODAGVYSARYIGNLFTSAFTLIVRCEAKQWGPCNHLCTACMNGVCHEDTGEC 240
DB 181 APODAGVYSARYIGNLFTSAFTLIVRCEAKQWGPCNHLCTACMNGVCHEDTGEC 240

QY 241 ICPPGFMGRTCKACELHTFGRTCKERCSGQGCKSYVFCPLDPYGCSCATCWGGLQNE 300
DB 241 ICPPGFMGRTCKACEPTFGRTCKERCSPGCKSYVFCPLDPYGCSCATCWGGLQNE 300

QY 301 ACHPGFYGDCKLRCSNNGENCDRFGQCLCSGQWGLQCEGREGIPRMTPKVILPDHIE 360
DB 301 ACPGSGYGYDCKLRCHCTNEECDFRFGQCLCSGQWGLQCEGREGIPRMTPKVILPDHIE 360

QY 361 VNSGKFNPKCKASGWLPLNEBMTLVKPDGTVLHFKDFTNTHFSAFTIHRILPPDSG 420
DB 361 VNSGKFNPKCKASGWLPLNEBMTLVKPDGTVLHFKDFTNTHFSAFTIHRILPPDSG 420

QY 421 VVWCSVNTVAGWVEKPFNISVKVLPKPLNPNVIDTGHNFVAINISSBYPFGDGPIKSKK 480
DB 421 VVWCSVNTVAGWVEKPFNISVKVLPKPLNPNVIDTGHNFVAINISSBYPFGDGPIKSKK 480

QY 481 LLYKPNHYEAWQHIOVTNEIVTLNLEPRTEYELCVQLVRRGEGEGHPGVRPRTTAS 540
DB 481 LLYKPNHYEAWQHIOVTNEIVTLNLEPRTEYELCVQLVRRGEGEGHPGVRPRTTAS 540

QY 541 IGLPPRGLNLLPKSGTTLNLTWQIPFSSBEDDFVEVERRSVQ-KSDQONIKVPGNLT 599
DB 539 IGLPPRGLNLLPKSGTTLNLTWQIPFSSBEDDFVEVERRSVQ-KSDQONIKVPGNLT 598

QY 600 VLLNNLHPREQVVRARVNTKAGWSEDLTAWTLDLPPQPMKISNITHSAVLSW 659
DB 599 VLLNNLHPREQVVRARVNTKAGWSEDLTAWTLDLPPQPMKISNITHSAVLSW 658

QY 660 TILDGYSISITIRYKVGKNEDQHVVDVKIKNATIIQYQLKLEPETAYQVDFIANNIG 719
DB 659 TIVDGYSSIIIRYKVGKNEDQHVVDVKIKNATIIQYQLKLEPETAYQVDFIANNIG 718

QY 720 SNNAPSHLVLTPESQAPADIGGKMLLIILGSGAGTCLTVLAFILQIKRANVQR 779
DB 719 SNNAPSHLVLTPESQADIGGKMLLIILGSGAGTCLTVLAFILQIKRANVQR 778

QY 780 RMAQAFQNVREBPVQFNGSTLALNRKVNNDPTIYPVLDWMDIKFQDVIGEGFGQVL 839
DB 779 RMAQAFQNVREBPVQFNGSTLALNRKVNNDPTIYPVLDWMDIKFQDVIGEGFGQVL 837

;
; REGISTRATION NUMBER: 34,971
; REFERENCE/DOCKET NUMBER: 3153-111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1122 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-838-957A-6

Query Match 92.7%; Score 5575; DB 1; Length 1122;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1041; Conservative 37; Mismatches 43; Indels 4; Gaps 3;

QY 840 KARIKKOGLRMDAAIKRMKEYASKDDHRDPAGELEVLCKLGHHPNIIINLGACEHGVLY 899
DB 838 KARIKKOGLRMDAAIKRMKEYASKDDHRDPAGELEVLCKLGHHPNIIINLGACEHGVLY 897

QY 900 LAIEYAPHGNLLDFLRKSRVLETPAFAIANSTASTLSSQQLLHFAADVARGMDYLSQKQ 959
DB 898 LAIEYAPHGNLLDFLRKSRVLETPAFAIANSTASTLSSQQLLHFAADVARGMDYLSQKQ 957

QY 960 FIHRDLAARNILVGENYVAKIADFGLSRGQEVYVKTWGRPLPVRWMAIESLNSYVYTNS 1019
DB 958 FIHRDLAARNILVGENYAKIADFGLSRGQEVYVKTWGRPLPVRWMAIESLNSYVYTNS 1017

QY 1020 DVWSYGVLLMEIVSLGGTPYCGMTCAELEYEKLPGQYRLEKPLNCDDVYVDMRCQWREK 1079
DB 1018 DVWSYGVLLMEIVSLGGTPYCGMTCAELEYEKLPGQYRLEKPLNCDDVYVDMRCQWREK 1077

QY 1080 YERPFAQILVSLNMLBERKTYNTTILYKFTYAGIDCSAEAA 1124
DB 1078 YERPFAQILVSLNMLBERKTYNTTILYKFTYAGIDCSAEAA 1122

RESULT 6
US-08-838-957A-6
; Sequence 6, Application US/08838957A
; Patent No. 5998187
; GENERAL INFORMATION:
; APPLICANT: Breitman, Martin L.
; APPLICANT: Rossant, Janet
; APPLICANT: Dumont, Daniel J.
; APPLICANT: Yamaguchi, Terry P.
; TITLE OF INVENTION: No. 5998187el Receptor Tyrosine Kinase
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,957A
; FILING DATE: 23-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurdydyk, Linda M.
; REGISTRATION NUMBER: 34,971
; REFERENCE/DOCKET NUMBER: 3153-212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1122 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-838-957A-6

Query Match 92.7%; Score 5575; DB 1; Length 1122;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1041; Conservative 37; Mismatches 43; Indels 4; Gaps 3;

QY 1 MDSLASLVLCGVSLLLSGTVEGMDLILINSPLVSDAETSLTCTASGWRPHEPTIGRD 60
DB 1 MDSLAGLVLCGVSLLLYGVVEGMDLILINSPLVSDAETSLTCTASGWHPEPTIGRD 60

QY 61 FEALNNQHDPLEVTQDVTREWAKKVVWKKREKASKINGAYFCEGRVGRGEAIRITMKMRQ 120
DB 61 FEALNNQHDPLEVTQDVTREWAKKVVWKKREKASKINGAYFCEGRVGRGEAIRITMKMRQ 120

QY 121 QASFLPATLTMTVDKGDNNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180
DB 121 QASFLPATLTMTVDKGDNNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180

QY 181 APODAGVYSARYIGNLFTSAFTLIVRCEAKQWGPCNHLCTACMNGVCHEDTGEC 240
DB 181 APODAGVYSARYIGNLFTSAFTLIVRCEAKQWGPCNHLCTACMNGVCHEDTGEC 240

QY 241 ICPPGFMGRTCKACELHTFGRTCKERCSGQGCKSYVFCPLDPYGCSCATCWGGLQNE 300
DB 241 ICPPGFMGRTCKACEPTFGRTCKERCSPGCKSYVFCPLDPYGCSCATCWGGLQNE 300

QY 301 ACHPGFYGDCKLRCSNNGENCDRFGQCLCSGQWGLQCEGREGIPRMTPKVILPDHIE 360
DB 301 ACPGSGYGYDCKLRCHCTNEECDFRFGQCLCSGQWGLQCEGREGIPRMTPKVILPDHIE 360

QY 361 VNSGKFNPKCKASGWLPLNEBMTLVKPDGTVLHFKDFTNTHFSAFTIHRILPPDSG 420
DB 361 VNSGKFNPKCKASGWLPLNEBMTLVKPDGTVLHFKDFTNTHFSAFTIHRILPPDSG 420

QY 421 VVWCSVNTVAGWVEKPFNISVKVLPKPLNPNVIDTGHNFVAINISSBYPFGDGPIKSKK 480
DB 421 VVWCSVNTVAGWVEKPFNISVKVLPKPLNPNVIDTGHNFVAINISSBYPFGDGPIKSKK 480

QY 481 LLYKPNHYEAWQHIOVTNEIVTLNLEPRTEYELCVQLVRRGEGEGHPGVRPRTTAS 540
DB 481 LLYKPNHYEAWQHIOVTNEIVTLNLEPRTEYELCVQLVRRGEGEGHPGVRPRTTAS 540

QY 541 IGLPPRGLNLLPKSGTTLNLTWQIPFSSBEDDFVEVERRSVQ-KSDQONIKVPGNLT 599
DB 539 IGLPPRGLNLLPKSGTTLNLTWQIPFSSBEDDFVEVERRSVQ-KSDQONIKVPGNLT 598

QY 600 VLLNNLHPREQVVRARVNTKAGWSEDLTAWTLDLPPQPMKISNITHSAVLSW 659
DB 599 VLLNNLHPREQVVRARVNTKAGWSEDLTAWTLDLPPQPMKISNITHSAVLSW 658

QY 660 TILDGYSISITIRYKVGKNEDQHVVDVKIKNATIIQYQLKLEPETAYQVDFIANNIG 719
DB 659 TIVDGYSSIIIRYKVGKNEDQHVVDVKIKNATIIQYQLKLEPETAYQVDFIANNIG 718

QY 720 SNNAPSHLVLTPESQAPADIGGKMLLIILGSGAGTCLTVLAFILQIKRANVQR 779
DB 719 SNNAPSHLVLTPESQADIGGKMLLIILGSGAGTCLTVLAFILQIKRANVQR 778

QY 780 RMAQAFQNVREBPVQFNGSTLALNRKVNNDPTIYPVLDWMDIKFQDVIGEGFGQVL 839
DB 779 RMAQAFQNVREBPVQFNGSTLALNRKVNNDPTIYPVLDWMDIKFQDVIGEGFGQVL 837
```

Db 61 FEALMNQHDPLEVTQDVTRWAKVVMKREKASKINGAYFCEGRVGRQAIRIRTMKMRQ 120  
Qy 121 QASFLPATLTMTVDKGDNNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
Db 121 QASFLPATLTMTVDKGDNNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
Qy 181 AQPDAGVYSARYIGNIFTSFTRLIVRRCBAQKWGPCNHLCTACMNGVCHEDTGEC 240  
Db 181 AQPDAGVYSARYIGNIFTSFTRLIVRRCBAQKWGPCNHLCTACMNGVCHEDTGEC 240  
Qy 241 ICPPGFMGRTCEKACELHTFGRTCKERCSCGQCKSVYFCLPDPYGCSCATGWGLQCN 300  
Db 241 ICPPGFMGRTCEKACELHTFGRTCKERCSCGQCKSVYFCLPDPYGCSCATGWGLQCN 300  
Qy 301 ACHPGFYGDPCKLRCSNNGEMCDRFOGCLCSGQCKSVYFCLPDPYGCSCATGWGLQCN 360  
Db 301 ACHPGFYGDPCKLRCSNNGEMCDRFOGCLCSGQCKSVYFCLPDPYGCSCATGWGLQCN 360  
Qy 361 VNSGKFNPICKASGWLPTNEEMTLVKPDGTVLHPKDFNHTDHFVAIFTHRIPLPDGSG 420  
Db 361 VNSGKFNPICKASGWLPTNEEMTLVKPDGTVLHPKDFNHTDHFVAIFTHRIPLPDGSG 420  
Qy 421 VVWCSVNTVAGWVEKPFNISVKVLPKPLNAPNVIDTGHNFVAVINISSEBYPFGDPIKSKK 480  
Db 421 VVWCSVNTVAGWVEKPFNISVKVLPKPLNAPNVIDTGHNFVAVINISSEBYPFGDPIKSKK 480  
Qy 481 LLYKPNVHAEWOHLQVNTNEIYVTLNLPRTBEYELCVQLVRGEGEGHGPVRRFTTAS 540  
Db 481 LLYKPNVHAEWOHLQVNTNEIYVTLNLPRTBEYELCVQLVRGEGEGHGPVRRFTTAS 540  
Qy 541 IGLPPRGNLNLPKSGOTTINLTWQIPFSSSEDDFVEVERRSVQ-KSQOQNIKVPGNLTS 599  
Db 599 IGLPPRGNLNLPKSGOTTINLTWQIPFSSSEDDFVEVERRSVQ-KSQOQNIKVPGNLTS 599  
Qy 600 VLLNNLHPREQVYVVRVNTKAAQSWSEDLTAWLSDILPPQENIKISNIHSSAVISW 659  
Db 600 VLLNNLHPREQVYVVRVNTKAAQSWSEDLTAWLSDILPPQENIKISNIHSSAVISW 659  
Qy 659 TILDCYSISSITIRKVKQKNEQDQVVKIKNATIIQYQKLEPETAYQVDIFANNIG 719  
Db 659 TILDCYSISSITIRKVKQKNEQDQVVKIKNATIIQYQKLEPETAYQVDIFANNIG 719  
Qy 720 SNNPASFHELVTLPSSQADILGGKMLLIATILGSAGMTCLTVLLAFILIIQLKRNQVR 779  
Db 720 SNNPASFHELVTLPSSQADILGGKMLLIATILGSAGMTCLTVLLAFILIIQLKRNQVR 779  
Qy 780 RMAQAFQNVREPAVQFNSGTILALNRKVKNPDPTIYPVLDWNDIKFQDVIQEGNFGQVL 839  
Db 780 RMAQAFQNVREPAVQFNSGTILALNRKVKNPDPTIYPVLDWNDIKFQDVIQEGNFGQVL 839  
Qy 840 KARIKKDGLRMDAAIKRMKEYASKDDHRDFAGELEVLCKLGHHPNIINLLGACERHGYLY 899  
Db 840 KARIKKDGLRMDAAIKRMKEYASKDDHRDFAGELEVLCKLGHHPNIINLLGACERHGYLY 899  
Qy 900 LAIEYAPGNLILDFURKSRVLETPDPAFAIANSTASTLSSQQLLHFAADVARGMDVLSQK 959  
Db 900 LAIEYAPGNLILDFURKSRVLETPDPAFAIANSTASTLSSQQLLHFAADVARGMDVLSQK 959  
Qy 960 FTHRDLAARNILVGENYVAKIADFLGRGQEVYVVKTMGRPLPVRMWAIESLNYSYVTTNS 1019  
Db 960 FTHRDLAARNILVGENYVAKIADFLGRGQEVYVVKTMGRPLPVRMWAIESLNYSYVTTNS 1019  
Qy 1020 DWWSYGVLLWEIVSLGGTPYCGMTCAELYEKLPQGRYLEKPLNCDDDEVYDLMRQCWRK 1079  
Db 1020 DWWSYGVLLWEIVSLGGTPYCGMTCAELYEKLPQGRYLEKPLNCDDDEVYDLMRQCWRK 1079  
Qy 1080 YERPSPFAQILVSLNRLMRKERTYVNTTILYEKFTYAGIDCSABEAA 1124  
Db 1080 YERPSPFAQILVSLNRLMRKERTYVNTTILYEKFTYAGIDCSABEAA 1124

; Sequence 2, Application US/07934393B  
; Patent No. 5466596  
; GENERAL INFORMATION:  
; APPLICANT: BREITMAN, MARTIN L.  
; APPLICANT: DUMONT, DANIEL  
; APPLICANT: GRADWOHL, GERARD G.  
; TITLE OF INVENTION: TISSUE SPECIFIC TRANSCRIPTIONAL  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BERESKIN & PARR  
; STREET: 40 King Street West  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5H 3Y2  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/934.393B  
; FILING DATE: 25-AUG-1992  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kurdydyk, Linda M.  
; REGISTRATION NUMBER: 34,971  
; REFERENCE/DOCKET NUMBER: 3153-64  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 354-7311  
; TELEFAX: (416) 361-1398  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1118 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-934-393B-2  
Query Match 92.2%; Score 5542; DB 1; Length 1118;  
Best Local Similarity 92.2%; Pred. No. 0;  
Matches 1037; Conservative 37; Mismatches 43; Indels 8; Gaps 4;  
Qy 1 MDSLASLVLCGVSLLSGTVEGAMDILLINSLPIVSDAETSLTCTASGWRPHEPTIGRD 60  
Db 1 MDSLAGLVLCGVSLLYGVGVEGAMDILLINSLPLVSDAETSLTCTASGWRPHEPTIGRD 60  
Qy 61 FEALMNQHDPLEVTQDVTRWAKVVMKREKASKINGAYFCEGRVGRQAIRIRTMKMRQ 120  
Db 61 FEALMNQHDPLEVTQDVTRWAKVVMKREKASKINGAYFCEGRVGRQAIRIRTMKMRQ 120  
Qy 121 QASFLPATLTMTVDKGDNNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
Db 121 QASFLPATLTMTVDKGDNNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
Qy 181 AQPDAGVYSARYIGNIFTSFTRLIVRRCBAQKWGPCNHLCTACMNGVCHEDTGEC 240  
Db 181 AQPDAGVYSARYIGNIFTSFTRLIVRRCBAQKWGPCNHLCTACMNGVCHEDTGEC 240  
Qy 241 ICPPGFMGRTCEKACELHTFGRTCKERCSCGQCKSVYFCLPDPYGCSCATGWGLQCN 300  
Db 241 ICPPGFMGRTCEKACELHTFGRTCKERCSCGQCKSVYFCLPDPYGCSCATGWGLQCN 300  
Qy 301 ACHPGFYGDPCKLRCSNNGEMCDRFOGCLCSGQCKSVYFCLPDPYGCSCATGWGLQCN 360  
Db 301 ACHPGFYGDPCKLRCSNNGEMCDRFOGCLCSGQCKSVYFCLPDPYGCSCATGWGLQCN 360  
Qy 361 VNSGKFNPICKASGWLPTNEEMTLVKPDGTVLHPKDFNHTDHFVAIFTHRIPLPDGSG 420  
Db 361 VNSGKFNPICKASGWLPTNEEMTLVKPDGTVLHPKDFNHTDHFVAIFTHRIPLPDGSG 420  
Qy 421 VVWCSVNTVAGWVEKPFNISVKVLPKPLNAPNVIDTGHNFVAVINISSEBYPFGDPIKSKK 480

Db 421 VVCSVNTVAGMVEKPFNISVKVLPPELHAPNVIDTGHNFPAINTSSEPYFGDPIKSKK 480  
Qy 481 LLYKPNVHYEAWOHIOVNEIIVTLNLEPRTEYELCVQVRGEGEGHPGVPVRRRTTAS 540  
Db 481 LFYKPNV--QAWKYIEVTNEIIFTLNLEPRTEYELCVQVRGEGEGHPGVPVRRRTTAC 538  
Qy 541 IGLPPRGNLPLPKSOTTLNLTWQPIFPSSSEDDFYVEVERRSVQ-KSDQONIKVPGNLTAS 599  
Db 539 IGLPPRGSLPLPKSOTALNLTWQPIFTNSEDEFYVEVERRSVQKSDQONIKVPGNLTAS 598  
Qy 600 VLLNLHPREQVYVVRARVNTKAGWSEDLTAWLTSLDILPPQENIKISNITHSSAVISW 659  
Db 599 VLLSNLVPREQYTVARVNTKAGWSEBELRAWLTSLDILPPQENIKISNITDSTAMVSW 658  
Qy 660 TILDGYSISITIRYKVGQKNEHDQVVKIKNATIIQYQKLGLEPETAYQVDFIAENNIG 719  
Db 659 TIVDGYSSIIIRYKVGQKNEHDQVVKIKNATTVQYQKLGLEPETTHVDFIAENNIG 718  
Qy 720 SNNPAPFSELVTLPEAQPADLGGGKMLLIALIGSAGMTCLTVLLAFLLIQLKRVANVOR 779  
Db 719 SNNPAPFSELRLTLPSPGADLGGGKMLLIALIGSAGMTCTITVLLAFLLIQLKRVANVOR 778  
Qy 780 RMAQAFONVREPAVQFNSGTLALNRKVNPNPTIYPVLDWNDIKFQDVIGEGNPGQVL 839  
Db 779 RMAQAFON-REPAVQFNSGTLALNRKAKNPNPTIYPVLDWNDIK----IGEGNPGQVL 833  
Qy 840 KARIKKGRLMDAAIKRMKEYASKDHRDPAGBLEVCKLGHHPNIIINLIGACEHRYLY 899  
Db 834 KARIKKGRLMDAAIKRMKEYASKDHRDPAGBLEVCKLGHHPNIIINLIGACEHRYLY 893  
Qy 900 LAIEYAPHGNLLDFLRKSRVLETPDPAFANSTASTLSSQQLLHFAADVARGMDYLSQKQ 959  
Db 894 LAIEYAPHGNLLDFLRKSRVLETPDPAFANSTASTLSSQQLLHFAADVARGMDYLSQKQ 953  
Qy 960 FHRDLAARNILVGENYAKIADFGLSRGQEVYVKTGMRLPVRWMAIESLNSYVYTTNS 1019  
Db 954 FHRDLAARNILVGENYAKIADFGLSRGQEVYVKTGMRLPVRWMAIESLNSYVYTTNS 1013  
Qy 1020 DWMSYGVLLWEIVSLGGTPYCGMTCAELYEKLPQGYRLEKPLNCDDEVYDLMRQWREKP 1079  
Db 1014 DWMSYGVLLWEIVSLGGTPYCGMTCAELYEKLPQGYRLEKPLNCDDEVYDLMRQWREKP 1073  
Qy 1080 YERPSPAQLVSLNRLMREKTYVNTTLYEKFTYAGIDCSABEAA 1124  
Db 1074 YERPSPAQLVSLNRLMREKTYVNTTLYEKFTYAGIDCSABEAA 1118

RESULT 8

US-08-278-089A-2  
; Sequence 2, Application US/08278089A  
; Patent No. 5681714  
; GENERAL INFORMATION:  
; APPLICANT: Breitman, Martin L.  
; APPLICANT: Rossant, Janet  
; APPLICANT: Dumont, Daniel J.  
; APPLICANT: Yamaguchi, Terry P.  
; TITLE OF INVENTION: No. 5681714el Receptor Tyrosine Kinase  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Bereskin & Parr  
; STREET: 40 King Street West  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5H 3Y2  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/278,089A

; FILING DATE: 20-JUL-1994  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kurdydyk, Linda M.  
; REGISTRATION NUMBER: 34,971  
; REFERENCE/DOCKET NUMBER: 3153-111  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 364-7311  
; TELEFAX: (416) 361-1398  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1118 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-278-089A-2  
  
Query Match 92.2%; Score 5542; DB 1; Length 1118;  
Best Local Similarity 92.2%; Pred. No. 0;  
Matches 1037; Conservative 37; Mismatches 43; Indels 8; Gaps 4;  
  
Qy 1 MDSLASLVLCGVSLLSGTVEGMDLILINSIPLVSDAETSLTCIASGWRPHEPIITGRD 60  
Db 1 MDSLALVLCGVSLLLYGVEGMDLILINSIPLVSDAETSLTCIASGWHPEPIITGRD 60  
Qy 61 FEALMNQHDQPLEVTQDVTREWAKKVVWREKASKINGAYFCEGRVGRGAIIRITMKMRQ 120  
Db 61 FEALMNQHDQPLEVTQDVTREWAKKVVWREKASKINGAYFCEGRVGRGAIIRITMKMRQ 120  
Qy 121 QASFLPATLTMTVDGDNVNISFKVLKKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
Db 121 QASFLPATLTMTVDGDNVNISFKVLKKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
Qy 181 APOQAGVYSARYIGNLFTSAFTRLIVRRCBAQKMGPECNHLCCTACMNGVCHEDTGE 240  
Db 181 APOQAGVYSARYIGNLFTSAFTRLIVRRCBAQKMGPECNHLCCTACMNGVCHEDTGE 240  
Qy 241 ICPPGFMGRTCEKACELHTFGRTCKERCQSGCKSYVFCCLPDPYGCSCATGWGLQCN 300  
Db 241 ICPPGFMGRTCEKACEPTFGRTCKERCQSGCKSYVFCCLPDPYGCSCATGWGLQCN 300  
Qy 301 ACHPGFYGDCKLRKSCNNGENCDRFOGCLCSGQGLQCEGREGIPRMTPKIVLDLPHIE 360  
Db 301 ACPGSGYGPDKLRKCHCTNEECDFRFOGCLCSGQGLQCEGREGIPRMTPKIVLDLPHIE 360  
Qy 361 VNSGKFNPICKASGWLPTNEBMTLVKPDGTVLHPKDFNHTDHFVIAFTIHRILPDDSG 420  
Db 361 VNSGKFNPICKASGWLPTNEBMTLVKPDGTVLHPKDFNHTDHFVIAFTIHRILPDDSG 420  
Qy 421 VVCSVNTVAGMVEKPFNISVKVLPKLPNAPNVIDTGHNFPAVINISSEPYFGDPIKSKK 480  
Db 421 VVCSVNTVAGMVEKPFNISVKVLPPELHAPNVIDTGHNFPAVINISSEPYFGDPIKSKK 480  
Qy 481 LLYKPNVHYEAWOHIOVNEIIVTLNLEPRTEYELCVQVRGEGEGHPGVPVRRRTTAS 540  
Db 481 LFYKPNV--QAWKYIEVTNEIIFTLNLEPRTEYELCVQVRGEGEGHPGVPVRRRTTAC 538  
Qy 541 IGLPPRGNLPLPKSOTTLNLTWQPIFPSSSEDDFYVEVERRSVQ-KSDQONIKVPGNLTAS 599  
Db 539 IGLPPRGSLPLPKSOTALNLTWQPIFTNSEDEFYVEVERRSVQKSDQONIKVPGNLTAS 598  
Qy 600 VLLNLHPREQVYVVRARVNTKAGWSEDLTAWLTSLDILPPQENIKISNITHSSAVISW 659  
Db 599 VLLSNLVPREQYTVARVNTKAGWSEBELRAWLTSLDILPPQENIKISNITDSTAMVSW 658  
Qy 660 TILDGYSISITIRYKVGQKNEHDQVVKIKNATIIQYQKLGLEPETAYQVDFIAENNIG 719  
Db 659 TIVDGYSSIIIRYKVGQKNEHDQVVKIKNATTVQYQKLGLEPETTHVDFIAENNIG 718  
Qy 720 SNNPAPFSELVTLPEAQPADLGGGKMLLIALIGSAGMTCLTVLLAFLLIQLKRVANVOR 779  
Db 719 SNNPAPFSELRLTLPSPGADLGGGKMLLIALIGSAGMTCTITVLLAFLLIQLKRVANVOR 778





RESULT 10

US-08-469-537A-96  
; Sequence 96, Application US/08469537A  
; Patent No. 5843749  
; GENERAL INFORMATION:  
; APPLICANT: Maisonnier, et al.  
; TITLE OF INVENTION: EHK AND ROR TYROSINE  
; TITLE OF INVENTION: KINASES  
; NUMBER OF SEQUENCES: 107  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
; STREET: 777 Old Saw Mill River Road  
; CITY: Tarrytown  
; STATE: NY  
; COUNTRY: U.S.A.  
; ZIP: 10591

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/469,537A  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: USSN 08/406,247  
; FILING DATE: 17-MAR-1995  
; APPLICATION NUMBER: USSN 08/144,992  
; FILING DATE: 28-OCT-1993  
; APPLICATION NUMBER: USSN 07/736,559  
; FILING DATE: 26-JUL-1991

ATTORNEY/AGENT INFORMATION:  
; NAME: Kempler, Ph.D., Gail M  
; REGISTRATION NUMBER: 32,143  
; REFERENCE/DOCKET NUMBER: REG 070C  
; TELEPHONE: 914-345-7400  
; TELEFAX: 914-345-7721  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 96:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1101 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-469-537A-96

Query Match 91.8%; Score 5518; DB 1; Length 1101;  
Best Local Similarity 93.3%; Pred. No. 0;  
Matches 1032; Conservative 30; Mismatches 38; Indels 6; Gaps 3;

QY	20	VEGAMDLILINSLPLVSDAETSLTCTIAGWRPHEPITIGRDPEALMNOHQDPLEVTQDVT	79
DB	1	VEGAMDLILINSLPLVSDAETSLTCTIAGWRPHEPITIGRDPEALMNOHQDPLEVTQDVT	60
QY	80	REWAKVWVKREKASKINGAYCEGRVGEATIRITMKRQOASFLPATLTMTVDKGNV	139
DB	61	REWAKVWVKREKASKINGAYCEGRVGEATIRITMKRQOASFLPATLTMTVDKGNV	120
QY	140	NISFKKVLKEEDAVTYKNGSIHSPRHEVPDILEVHLPHAQPODAGVYSARYIGGNLF	199
DB	121	NISFKKVLKEEDAVTYKNGSIHSPRHEVPDILEVHLPHAQPODAGVYSARYIGGNLF	180
QY	200	TSAFTRLIVRRCEAKWGPCNHLCTACMNGVCHEDTGECICPFGMGRTCCEKACELHT	259
DB	181	TSAFTRLIVRRCEAKWGPCNHLCTACMNGVCHEDTGECICPFGMGRTCCEKACEPHT	240
QY	260	FORTCKERCSSGEGCKSVYFCLPDPYGCSCATGWGLQCNBACHPGFVGPDPCKLRCSNN	319

DB	241	FORTCKERCSSGEGCKSVYFCLPDPYGCSCATGWGLQCNBACHPGFVGPDPCKLRCHCTN	300
QY	320	GEMCDRFQGCCLSPGWQGLQCREGIPRMTPKIVDLPDHIENVSCKFNPKICKASGWPLPT	379
DB	301	EEMCDRFQGCCLSPGWQGLQCEKEGRPMTPQIEDLPDHIENVSCKFNPKICKASGWPLPT	360
QY	380	NEEMTLVKPDGTVLHPKOPNHTDHFSAIFTHRIILPDSGVVWCVSVNTVAGWVEKPFNI	439
DB	361	SEEMTLVKPDGTVLQPNDFNHTDHFSAIFTVNRILPDSGVVWCVSVNTVAGWVEKPFNI	420
QY	440	SVKVLKPLNAPNVIDTGNFAVINISSEPYGDCGPIKSKLLYKPVNHYEAWOHIQVTN	499
DB	421	SVKVLKPLNAPNVIDTGNFAVINISSEPYGDCGPIKSKLLYKPVNHYEAWOHIQVTN	478
QY	500	EIVTLNLYLEPRTEYELCVLVRGEGGEGHPVRRFTTASIGLPPRGILNLLPKSQTL	559
DB	479	EIVTLNLYLEPRTDYELCVLVRGEGGEGHPVRRFTTASIGLPPRGILNLLPKSQTL	538
QY	560	NLTWQPIFPSSDDPFYVEVERSVQ-KSDQONIKVPGNLTSLVLLNNLHPRQYVVRVYN	618
DB	539	NLTWQPIFTSSEDEFYVEVERWSQTRSDQQNKKVPGNLTSLVLLNNLHPRQYVVRVYN	598
QY	619	TKAQGEWSEDLTAWTLDLPPQENIKISNTHSSAVISWTILDGYSISSITIRYKVOG	678
DB	599	TKAQGEWSEELRATLSDILPPQENIKITNTDTALVSWITVDGYSISSIIIRYKVOG	658
QY	679	KNEDQHVVDVKKNATIIQVQLKGLPETAYQVDIIPAENNIGSSNPAFSLHVLTPESQAP	738
DB	659	KNEDQHIDVKKNATITQVQLKGLPETTYHYDIPAENNIGSSNPAFSEQEITLP--AP	715
QY	739	ADLGGKMLLITAILGSAGMTCTVLLAFILILQKRVNQRMAQAFQVVRREPAPQVNS	798
DB	716	KDLGGKMLLITAILGSAGMTCTVLLAFILIMLQKRVNQRMAQAFQVVRREPAPQVNS	775
QY	799	GTALNRKVNKNPDPTIYVLDNDNDIKFQDVTGEGNFGQVLRKIKDGLRMDAAIKRMK	858
DB	776	GTALNRKAKNPDPTIYVLDNDNDIKFQDVTGEGNFGQVLRKIKDGLRMDAAIKRMK	835
QY	859	EYASKDDHRDFAGELEVLCKLGHHPNIIINLLGACHEHGYLYLAIEYAPHGNLLDFLRKR	918
DB	836	EYASKDDHRDFAGELEVLCKLGHHPNIIINLLGACHEHGYLYLAIEYAPHGNLLDFLRKR	895
QY	919	VLETPAFAIANSTASTLSQQLLHFAADVARGMDYLSQKQFIHRDLAARNLLVGENYVA	978
DB	896	VLETPAFAIANSTASTLSQQLLHFAADVARGMDYLSQKQFIHRDLAARNLLVGENYVA	955
QY	979	KIADFGLSRGQEVYVKTGMRLPVRWMALESINYSVYTTNSDVWSVGLLWEIVSIGGTP	1038
DB	956	KIADFGLSRGQEVYVKTGMRLPVRWMALESINYSVYTTNSDVWSVGLLWEIVSIGGTP	1015
QY	1039	YCGMTCAELYEKLPGYRLEKPLNCDDEYDLMRQCWREKPYERPSFAQILVSLNRLME	1098
DB	1016	YCGMTCAELYEKLPGYRLEKPLNCDDEYDLMRQCWREKPYERPSFAQILVSLNRLME	1075
QY	1099	RKYTYNTLYEKFYTAGIDCSABEAA	1124
DB	1076	RKYTYNTLYEKFYTAGIDCSABEAA	1101

RESULT 11

US-09-590-656-1  
; Sequence 1, Application US/09590656  
; Patent No. 6413932  
; GENERAL INFORMATION:  
; APPLICANT: Cerretti, Douglas P.  
; APPLICANT: Borges, Luis G.  
; APPLICANT: Fanslow, III, William C.  
; TITLE OF INVENTION: TEK ANTAGONISTS  
; FILE REFERENCE: 2900-A  
; CURRENT APPLICATION NUMBER: US/09/590,656  
; CURRENT FILING DATE: 2000-06-07  
; PRIOR APPLICATION NUMBER: 60/137,889  
; PRIOR FILING DATE: 1999-06-07



Qy	601	LLNNLHPREQYVVRARVNTKAQGWS	EDLTATWTLSDILPPQPNIKISNITHSSA	VISMT	660
Dd	601	LLNNLHPREQYVVRARVNTKAQGWS	EDLTATWTLSDILPPQPNIKISNITHSSA	VISMT	660
Qy	661	ILDGYSIISITIRYKVGKNEDQHVDVK	KINATIIYOQLKGLSPETAYQVDIFAEN	NIGS	720
Dd	661	ILDGYSIISITIRYKVGKNEDQHVDVK	KINATIIYOQLKGLSPETAYQVDIFAEN	NIGS	720
Qy	721	SNPASFSELVTLPESQA	PADLGCGK-----MLL-----	----	748
Dd	721	SNPASFSELVTLPESQA	PADLGCGK-----MLL-----	----	748
Qy	749	IALLSGAGMTCUTVLLAFILILQKL	RANVRORMAQAFONVREBP	AQVFN---SGTIALNR	805
Dd	781	LMISRTEPTCVTVDDVS-----	HEDPEVKFNMVYDGVGVHNA	817	
Qy	806	KVKNNPDP--TIYPVL	DNDIKFDIVIGEGFGQVLXARIKKOGL--RMDAAIKRMKEVA	861	
Dd	818	KTKPREEQNSTYRVSVLVTL	VHODWLN----GKEYCKVSNKALPAPIEKTISKAKGP	873	
Qy	862	SXODHRDFAGELEVLCKLGHIPNI	INILGACEHRGYL--YLATEYAPHGN-LLDPLRKSR	918	
Dd	874	REPQVYTLPSPREEMTK-----	NQVSL--TCLVKGFPSDI	AVESNGQPENNYKTTTP	926
Qy	919	VLETDPAP	926		
Dd	927	VLDSDGSP	934		
RESULT 13					
US-09-590-656-2					
; Sequence 2, Application US/09590656					
; Patent No. 6413932					
; GENERAL INFORMATION:					
; APPLICANT: Cerretti, Douglas P.					
; APPLICANT: Borges, Luis G.					
; APPLICANT: Fanslow, III, William C.					
; TITLE OF INVENTION: TEK ANTAGONISTS					
; FILE REFERENCE: 2900-A					
; CURRENT APPLICATION NUMBER: US/09/590,656					
; CURRENT FILING DATE: 2000-06-07					
; PRIOR APPLICATION NUMBER: 60/137,889					
; PRIOR FILING DATE: 1999-06-07					
; NUMBER OF SEQ ID NOS: 2					
; SOFTWARE: PatentIn Ver. 2.0					
; SEQ ID NO 2					
; LENGTH: 704					
; TYPE: PRT					
; ORGANISM: Homo sapiens					
US-09-590-656-2					
Query Match 43.6%; Score 2623; DB 2; Length 704;					
Best Local Similarity 77.1%; Pred. No. 4.5e-161;					
Matches 501; Conservative 23; Mismatches 44; Indels 82; Gaps 7					
Qy	1	MDSLASLVLCGVSILLSGTVEGAMD	LILNSLPLVSDAETS LTCIASGWRPHEP	ITIGRD	60
Dd	1	MDSLASLVLCGVSILLSGTVEGAMD	LILNSLPLVSDAETS LTCIASGWRPHEP	ITIGRD	60
Qy	61	FEALMNHQHDPLEVTDV	TREMAKVVVMKREKASKINGAYFCSEGR	VRGEAIRIRTMMKQ	120
Dd	61	FEALMNHQHDPLEVTDV	TREMAKVVVMKREKASKINGAYFCSEGR	VRGEAIRIRTMMKQ	120
Qy	121	QASFLPATLTMTVDKGDNVN	ISFKKVLIKEDDAVIYKNGSF	FHSVRPHEVPDILEVHLPH	180
Dd	121	QASFLPATLTMTVDKGDNVN	ISFKKVLIKEDDAVIYKNGSF	FHSVRPHEVPDILEVHLPH	180
Qy	181	APOQDAGVSARYIGNLI	FTSAFTRLIVRRCEAKWGPECNHLCTAC	NNNGVCHEDTGRC	240
Dd	181	APOQDAGVSARYIGNLI	FTSAFTRLIVRRCEAKWGPECNHLCTAC	NNNGVCHEDTGRC	240
Qy	241	ICPPGFMGRTCBACELHT	FGRTCKERGSGOGCKSVFCLPD	PGYGSCATGWKGLQCNE	300





“This Page Blank (uspto)”

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2006, 01:59:33 ; Search time 177 Seconds  
(without alignments)  
2653.334 Million cell updates/sec

Title: US-10-633-742-8

Perfect score: 6013

Sequence: 1 MDLSASLVLCVSGVLLSGTV.....TTLYEKFTYAGIDCSAERAA 1124

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA\_Main:\*

- 1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 3: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 4: /cgn2\_6/prodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 5: /cgn2\_6/prodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 6: /cgn2\_6/prodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6013	100.0	1124	4	US-10-223-085-82
2	6013	100.0	1124	4	US-10-223-084-82
3	6013	100.0	1124	4	US-10-223-088-82
4	6013	100.0	1124	4	US-10-223-090-82
5	6013	100.0	1124	4	US-10-223-087-82
6	6013	100.0	1124	4	US-10-223-083-82
7	6013	100.0	1124	4	US-10-223-089-82
8	6013	100.0	1124	4	US-10-021-660-112
9	6013	100.0	1124	4	US-10-223-081-82
10	6013	100.0	1124	4	US-10-223-082-82
11	6013	100.0	1124	4	US-10-305-654-82
12	6013	100.0	1124	4	US-10-394-322A-64
13	6013	100.0	1124	4	US-10-211-462-147
14	6013	100.0	1124	4	US-10-081-056-82
15	6013	100.0	1124	6	US-11-019-855-47
16	6013	100.0	1127	3	US-09-815-341A-1
17	5732.5	95.3	1125	5	US-10-732-923-13691
18	5580	92.8	1122	5	US-10-732-923-13434
19	5413	90.0	1083	5	US-10-732-923-13449
20	4612.5	76.7	864	5	US-10-846-113A-24
21	4311	71.7	821	5	US-10-846-113A-23
22	4203	69.9	798	5	US-10-846-113A-22
23	4043	67.2	977	3	US-09-733-764-1
24	4043	67.2	977	4	US-10-357-653-1
25	2652	44.1	949	4	US-10-232-838-19
26	2639.5	43.9	484	4	US-10-232-838-30
27	2623	43.6	704	3	US-09-733-764-2

28	2623	43.6	704	4	US-10-357-653-2	Sequence 2, Appli
29	2606	43.3	810	4	US-10-232-838-15	Sequence 15, Appl
30	2567	42.7	949	4	US-10-232-838-18	Sequence 18, Appl
31	2550.5	42.4	465	4	US-10-232-838-29	Sequence 29, Appl
32	2544	42.3	934	4	US-10-232-838-17	Sequence 17, Appl
33	2527	42.0	1138	4	US-10-021-660-101	Sequence 101, App
34	2527	42.0	1138	4	US-10-394-322A-63	Sequence 63, Appl
35	2527	42.0	1138	4	US-10-211-462-113	Sequence 113, App
36	2527	42.0	1138	5	US-10-916-840-2	Sequence 2, Appli
37	2527	42.0	1138	5	US-10-846-113A-28	Sequence 28, Appl
38	2524	42.0	454	4	US-10-232-838-28	Sequence 28, Appl
39	2451	41.8	1134	4	US-10-116-275-137	Sequence 137, App
40	2451	40.8	658	5	US-10-732-923-13643	Sequence 13643, A
41	1847	30.7	351	4	US-10-664-421-58	Sequence 58, Appl
42	1847	30.7	351	5	US-10-941-635-58	Sequence 58, Appl
43	1791	29.8	736	3	US-10-450-763-45543	Sequence 45543, A
44	1703	28.3	326	3	US-09-815-341A-2	Sequence 2, Appli
45	1680	27.9	327	4	US-10-664-421-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1

US-10-223-085-82  
; Sequence 82, Application US/10223085  
; Publication NO. US20030100497A1  
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Marsters, Scot A.  
APPLICANT: Pan, James  
APPLICANT: Stephan, Jean-Philippe F.  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Williams, P.Mickey  
APPLICANT: Ye, Weilan  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF DISORDERS INVOLVING ANGIOGENESIS  
FILE REFERENCE: P3235PIC10  
CURRENT FILING DATE: 2002-08-16  
PRIOR FILING DATE: 2002-03-20  
PRIOR FILING DATE: 2002-03-20  
PRIOR FILING DATE: 2000-06-23  
PRIOR FILING DATE: 2000-07-20  
PRIOR FILING DATE: 2000-07-20  
PRIOR FILING DATE: 2000-07-25  
PRIOR FILING DATE: 2000-07-25  
PRIOR FILING DATE: 2000-07-28  
PRIOR FILING DATE: 2000-08-02  
PRIOR FILING DATE: 2000-08-17  
PRIOR FILING DATE: 2000-08-17  
PRIOR FILING DATE: 2000-08-23  
PRIOR FILING DATE: 2000-08-24  
NUMBER OF SEQ ID NOS: 383  
SEQ ID NO 82  
LENGTH: 1124  
TYPE: PRT

Db	961	IHRDLAARNILVGENYAKIADFGLSRGQEVYVVKTKMGRLPVRWMAIESLNSVYTTNSD	1020
Qy	1021	VMSYGVTLWEIVSLGCTPYCGMTCALYEKLPQGYRLEKPLNCDDDEVVDLMRQCWREKPY	1080
Db	1021	VMSYGVTLWEIVSLGCTPYCGMTCALYEKLPQGYRLEKPLNCDDDEVVDLMRQCWREKPY	1080
Qy	1081	ERPSAQILVSLNRMLEBKTYVNTTLYEKFTYAGIDCSAEAA	1124
Db	1081	ERPSAQILVSLNRMLEBKTYVNTTLYEKFTYAGIDCSAEAA	1124
RESULT 2			
US-10-223-084-82			
; Sequence 82, Application US/10223084			
; Publication No. US20030105011A1			
; GENERAL INFORMATION:			
; APPLICANT: Baker, Kevin P.			
; APPLICANT: Ferrara, Napoleone			
; APPLICANT: Gerber, Hanspeter			
; APPLICANT: Gerritsen, Mary E.			
; APPLICANT: Goddard, Audrey			
; APPLICANT: Godowski, Paul J.			
; APPLICANT: Gurney, Austin L.			
; APPLICANT: Hillan, Kenneth J.			
; APPLICANT: Masters, Scot A.			
; APPLICANT: Pan, James			
; APPLICANT: Stephan, Jean-Philippe P.			
; APPLICANT: Watanabe, Colin K.			
; APPLICANT: Wood, William I.			
; APPLICANT: Williams, P.Mickey			
; APPLICANT: Ye, Weilan			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND			
; TITLE OF INVENTION: TREATMENT OF DISORDERS INVOLVING ANGIOGENESIS			
; FILE REFERENCE: P3235PICS			
; CURRENT APPLICATION NUMBER: US/10/223,084			
; CURRENT FILING DATE: 2002-08-16			
; PRIOR APPLICATION NUMBER: US 10/081,056			
; PRIOR FILING DATE: 2002-02-20			
; PRIOR APPLICATION NUMBER: US 60/213,637			
; PRIOR FILING DATE: 2000-06-23			
; PRIOR APPLICATION NUMBER: US 60/219,556			
; PRIOR FILING DATE: 2000-07-20			
; PRIOR APPLICATION NUMBER: US 60/220,624			
; PRIOR FILING DATE: 2000-07-25			
; PRIOR APPLICATION NUMBER: US 60/220,664			
; PRIOR FILING DATE: 2000-07-25			
; PRIOR APPLICATION NUMBER: PCT/US00/20710			
; PRIOR FILING DATE: 2000-07-28			
; PRIOR APPLICATION NUMBER: US 60/222,695			
; PRIOR FILING DATE: 2000-08-02			
; PRIOR APPLICATION NUMBER: US 09/643,657			
; PRIOR FILING DATE: 2000-08-17			
; PRIOR APPLICATION NUMBER: PCT/US00/23522			
; PRIOR FILING DATE: 2000-08-23			
; PRIOR APPLICATION NUMBER: PCT/US00/23328			
; PRIOR FILING DATE: 2000-08-24			
; Remaining Prior Application data removed - See File Wrapper or PALM.			
; NUMBER OF SEQ ID NOS: 383			
; SEQ ID NO 82			
; LENGTH: 1124			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-223-084-82			
Query Match 100.0%; Score 6013; DB 4; Length 1124;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MDSLASVLGVSLLSGTVEGAMDILILINSIPLVSDAETSLTCTASGWRPHEPITIGRD	60
Db	1	MDSLASVLGVSLLSGTVEGAMDILILINSIPLVSDAETSLTCTASGWRPHEPITIGRD	60
Qy	61	FEALMNQHDPLEVTQDVTREWAKVVMKREKASKINGAYFCEGRVGEAIRITMKMRQ	120
Db	61	FEALMNQHDPLEVTQDVTREWAKVVMKREKASKINGAYFCEGRVGEAIRITMKMRQ	120
Qy	121	QASFLPATLTMTVDKGNVNISFKVLKEEDAVIYKNGSFTHSVPRHVPDILEVHLPH	180
Db	121	QASFLPATLTMTVDKGNVNISFKVLKEEDAVIYKNGSFTHSVPRHVPDILEVHLPH	180
Qy	181	AQPDAGVYSARYIGNLFTSAFTRLIVRCEAQKWGPCNHLCTACMNGVCHEDTSEC	240
Db	181	AQPDAGVYSARYIGNLFTSAFTRLIVRCEAQKWGPCNHLCTACMNGVCHEDTSEC	240
Qy	241	ICPPGFMGRTCKACELHTFGTRCKERCSCGQSGCKSYVFCPLDPYGCSCATGMKGLQNE	300
Db	241	ICPPGFMGRTCKACELHTFGTRCKERCSCGQSGCKSYVFCPLDPYGCSCATGMKGLQNE	300
Qy	301	ACHPGFYGDPCKLRCSCNNGCNDPFGCLSPGQGLQCEBEGIPRMTPKIVDLPDIE	360
Db	301	ACHPGFYGDPCKLRCSCNNGCNDPFGCLSPGQGLQCEBEGIPRMTPKIVDLPDIE	360
Qy	361	VNSGFNPKICKASGWLPTNEEMTLVKPDGTVLHPKDNHTDHFSAFTTHRIILPPDSG	420
Db	361	VNSGFNPKICKASGWLPTNEEMTLVKPDGTVLHPKDNHTDHFSAFTTHRIILPPDSG	420
Qy	421	VVWCSVNTVAGWVEKPFNISVKVLPKLPNAPNVIDTGNFAVINISSEPPYFGDGPISKK	480
Db	421	VVWCSVNTVAGWVEKPFNISVKVLPKLPNAPNVIDTGNFAVINISSEPPYFGDGPISKK	480
Qy	481	LLYKPNHYEAWQHIQVNTNEIYTLNLPRTYEYLCVQLVRGEGBEGHGPVRRFTAS	540
Db	481	LLYKPNHYEAWQHIQVNTNEIYTLNLPRTYEYLCVQLVRGEGBEGHGPVRRFTAS	540
Qy	541	IGLPPRGNLNLPKSOITLNTWQIPFSSSEDDFVVEVERRSVQSKDOQNIKVPGNLTSV	600
Db	541	IGLPPRGNLNLPKSOITLNTWQIPFSSSEDDFVVEVERRSVQSKDOQNIKVPGNLTSV	600
Qy	601	LNNLHPREYVVRARVNTKAGWSEDLTAWTSLDILPPQPENIKISNITHSSAVISWT	660
Db	601	LNNLHPREYVVRARVNTKAGWSEDLTAWTSLDILPPQPENIKISNITHSSAVISWT	660
Qy	661	ILDGYSISITIRYKVGKNEHDQVDVKIKNATIIQYQLKLEPETAYQVDIFAEENIGS	720
Db	661	ILDGYSISITIRYKVGKNEHDQVDVKIKNATIIQYQLKLEPETAYQVDIFAEENIGS	720
Qy	721	SNPAFSELVLPESQAPADLGGGKMLLIALGSGAGMTCLTVLLAFLLIQLKRVNQR	780
Db	721	SNPAFSELVLPESQAPADLGGGKMLLIALGSGAGMTCLTVLLAFLLIQLKRVNQR	780
Qy	781	MAQAFONVREBPAGVNSGTLALNRKVNPNDDTIYPVLDMDNDIKFQDVGEGNGQVLK	840
Db	781	MAQAFONVREBPAGVNSGTLALNRKVNPNDDTIYPVLDMDNDIKFQDVGEGNGQVLK	840
Qy	841	ARIKKGGLRMDAAIKRMKEYASKDHRDPAGEVLCKLGHHPNLIINLGACEHGGYLYL	900
Db	841	ARIKKGGLRMDAAIKRMKEYASKDHRDPAGEVLCKLGHHPNLIINLGACEHGGYLYL	900
Qy	901	ALEYAPHGNLLDFLRKSRVLETDPAFANSTASTLSSQQLLHFAADVARGMDYLSQKQF	960
Db	901	ALEYAPHGNLLDFLRKSRVLETDPAFANSTASTLSSQQLLHFAADVARGMDYLSQKQF	960
Qy	961	IHRDLAARNILVGENYAKIADFGLSRGQEVYVVKTKMGRLPVRWMAIESLNSVYTTNSD	1020



```
Db 61 FEALMNQHDPLEVTDVTRWAKVWVWREKASKINGAYFCEGRVGEAIRITMQRQ 120
Qy 121 QASFLPATLTMTVDKGDVNIISFKKVLKEEDAVIYKNGSFHISVPRHEVPDILEVHLPH 180
Db 121 QASFLPATLTMTVDKGDVNIISFKKVLKEEDAVIYKNGSFHISVPRHEVPDILEVHLPH 180
Qy 181 AQPQDAGVYSARYIGGNLFTSAFTRLIVRRCEAQKWGPCNHLCTACMNGVCHEDTGEC 240
Db 181 AQPQDAGVYSARYIGGNLFTSAFTRLIVRRCEAQKWGPCNHLCTACMNGVCHEDTGEC 240
Qy 241 ICPPGFMGRCTCKACELHFTFGRTCKERCSCGQCKSYVFCCLDPDYGCSCATWKGLQCNE 300
Db 241 ICPPGFMGRCTCKACELHFTFGRTCKERCSCGQCKSYVFCCLDPDYGCSCATWKGLQCNE 300
Qy 301 ACHPGFYGDCKLRCSNNGEMCDRFQGLCLSPGMOGLQCEGREGIPRMTPKIVDLPDHE 360
Db 301 ACHPGFYGDCKLRCSNNGEMCDRFQGLCLSPGMOGLQCEGREGIPRMTPKIVDLPDHE 360
Qy 361 VNSGKFNPICKASGWLPTNBEWTLVKPDGTVLHPKDFNHTDHFVAIPTIHRILPPDSG 420
Db 361 VNSGKFNPICKASGWLPTNBEWTLVKPDGTVLHPKDFNHTDHFVAIPTIHRILPPDSG 420
Qy 421 VNVCSVNTVAGVVEKPFNLSVKVLPKPLNAPNVIDTGHNFVNINISSEYPFGDGPIKSKK 480
Db 421 VNVCSVNTVAGVVEKPFNLSVKVLPKPLNAPNVIDTGHNFVNINISSEYPFGDGPIKSKK 480
Qy 481 LLYKPNHYEAWQHIOVTNEIIVTLNLEPRTYEYELCVQVLRGEGEGHPGVPRFTTAS 540
Db 481 LLYKPNHYEAWQHIOVTNEIIVTLNLEPRTYEYELCVQVLRGEGEGHPGVPRFTTAS 540
Qy 541 IGLPPRGNLNLPKSGQTLNLTPQIPFSSDDFYVEVERRSVQKSDQONIKVPGNLTSV 600
Db 541 IGLPPRGNLNLPKSGQTLNLTPQIPFSSDDFYVEVERRSVQKSDQONIKVPGNLTSV 600
Qy 601 LNNLHPROYVVRARVNTKAQGESEDLTAWTLDILPPQENIKISNITHSSAVISWT 660
Db 601 LNNLHPROYVVRARVNTKAQGESEDLTAWTLDILPPQENIKISNITHSSAVISWT 660
Qy 661 ILDGYSISITIRYKVGKNEQHDVVKIKNATIIOYQLKLEPETAYOVDFIFAENNIGS 720
Db 661 ILDGYSISITIRYKVGKNEQHDVVKIKNATIIOYQLKLEPETAYOVDFIFAENNIGS 720
Qy 721 SNPAFSELVTLPEQAPADLGGKXMLLAILGSGMTCLTVLLAFLIILQKLANVQR 780
Db 721 SNPAFSELVTLPEQAPADLGGKXMLLAILGSGMTCLTVLLAFLIILQKLANVQR 780
Qy 781 MAQAFQNVREEPAVQFNSGTLALNRKVKNPDPPTIYPVLDWNDIKFQDVIGEGFQVLK 840
Db 781 MAQAFQNVREEPAVQFNSGTLALNRKVKNPDPPTIYPVLDWNDIKFQDVIGEGFQVLK 840
Qy 841 ARIKXGDLRMDAAIKMKKEYASKDHRDPAGLEVLCKLGHHPNIIINLGACEHGYLYL 900
Db 841 ARIKXGDLRMDAAIKMKKEYASKDHRDPAGLEVLCKLGHHPNIIINLGACEHGYLYL 900
Qy 901 AIEYAPHGNLDFLRKSRVLETPDPAFANSTASTLSSQQLLHFAADVARGMDYLSQKF 960
Db 901 AIEYAPHGNLDFLRKSRVLETPDPAFANSTASTLSSQQLLHFAADVARGMDYLSQKF 960
Qy 961 IHRDLAARNILGVENVAKIADFGLSRGQEVVYKKTGMRLPVRWMAIESLNTSVYTTNSD 1020
Db 961 IHRDLAARNILGVENVAKIADFGLSRGQEVVYKKTGMRLPVRWMAIESLNTSVYTTNSD 1020
Qy 1021 VNSYGVLLWEIVSLGTPCYGTMTCABLYEKLPGQYRLEKPLNCDDEVDYDLMQCQREKPY 1080
Db 1021 VNSYGVLLWEIVSLGTPCYGTMTCABLYEKLPGQYRLEKPLNCDDEVDYDLMQCQREKPY 1080
Qy 1081 ERPSFAQILVSLNRMLEERKTYNTTLYEKFYTAGIDCSAEEAA 1124
Db 1081 ERPSFAQILVSLNRMLEERKTYNTTLYEKFYTAGIDCSAEEAA 1124
```

```
US-10-223-088-82
; Sequence 82, Application US/10223088
; Publication NO. US20030105012A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Williams, P.Mickey
; APPLICANT: Ye, Weilan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: P3235PIC6
; CURRENT APPLICATION NUMBER: US/10/223,088
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 10/081,056
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: US 60/213,637
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/219,556
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: US 60/220,624
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/220,664
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: PCT/US00/20710
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/222,695
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: US 09/643,657
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: PCT/US00/23522
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 383
; SEQ ID NO 82
; LENGTH: 1124
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-223-088-82
```

```
Query Match 100.0%; Score 6013; DB 4; Length 1124;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDSLASLVLCGVSLLSGTVEGAMDLIILNSIPLVSDAETSLTCTASGWRPHEPITIGRD 60
Db 1 MDSLASLVLCGVSLLSGTVEGAMDLIILNSIPLVSDAETSLTCTASGWRPHEPITIGRD 60
Qy 61 FEALMNQHDPLEVTDVTRWAKVWVWREKASKINGAYFCEGRVGEAIRITMQRQ 120
Db 61 FEALMNQHDPLEVTDVTRWAKVWVWREKASKINGAYFCEGRVGEAIRITMQRQ 120
Qy 121 QASFLPATLTMTVDKGDVNIISFKKVLKEEDAVIYKNGSFHISVPRHEVPDILEVHLPH 180
Db 121 QASFLPATLTMTVDKGDVNIISFKKVLKEEDAVIYKNGSFHISVPRHEVPDILEVHLPH 180
Qy 181 AQPQDAGVYSARYIGGNLFTSAFTRLIVRRCEAQKWGPCNHLCTACMNGVCHEDTGEC 240
Db 181 AQPQDAGVYSARYIGGNLFTSAFTRLIVRRCEAQKWGPCNHLCTACMNGVCHEDTGEC 240
Qy 241 ICPPGFMGRCTCKACELHFTFGRTCKERCSCGQCKSYVFCCLDPDYGCSCATWKGLQCNE 300
```

Db 241 ICPGFMGRTECEKACELHTFGRTCKERCSCGQCKSYVFCCLPDPYGCSCATGKGLQCNH 300  
Qy 301 ACHPGFYGPDCKLRCSCNNGEMCDRFOGCLSPGQGLQCEBEGIPRMTPKIVDLPDHE 360  
Db 301 ACHPGFYGPDCKLRCSCNNGEMCDRFOGCLSPGQGLQCEBEGIPRMTPKIVDLPDHE 360  
Qy 361 VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAVFTTHRIILPPDSG 420  
Db 361 VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAVFTTHRIILPPDSG 420  
Qy 421 VVCSVNTVAGWVEKPFNISVKVLPRKPLNAPNVIDTGHNFVAVINTSSBYPFGDGPIKSKK 480  
Db 421 VVCSVNTVAGWVEKPFNISVKVLPRKPLNAPNVIDTGHNFVAVINTSSBYPFGDGPIKSKK 480  
Qy 481 LLYKPNHYEAWOHIQVNTNEIYTLNLEPRTEYELCVQVLRGEGEGHGPVRFPTTAS 540  
Db 481 LLYKPNHYEAWOHIQVNTNEIYTLNLEPRTEYELCVQVLRGEGEGHGPVRFPTTAS 540  
Qy 541 IGLPPRGILNLLPKSOTTLNLTWQPIFPSSDDFVVEVRRSVQKSDQONIKVPGNLTSV 600  
Db 541 IGLPPRGILNLLPKSOTTLNLTWQPIFPSSDDFVVEVRRSVQKSDQONIKVPGNLTSV 600  
Qy 601 LNNLHPREQYVVRARVNTKAQWSEDLTAWTSLDILPPQENIKISNITHSSAVISWT 660  
Db 601 LNNLHPREQYVVRARVNTKAQWSEDLTAWTSLDILPPQENIKISNITHSSAVISWT 660  
Qy 661 ILDGYSISSITIRYKVGKNEQHVVDVKIKNATIIQYQKLEPETAYQVDIFAENNIGS 720  
Db 661 ILDGYSISSITIRYKVGKNEQHVVDVKIKNATIIQYQKLEPETAYQVDIFAENNIGS 720  
Qy 721 SNPAFSHELVTLPESQAPADLGGGMILLIILGSGAGMTCLTVLLAFLIILQKRVNORR 780  
Db 721 SNPAFSHELVTLPESQAPADLGGGMILLIILGSGAGMTCLTVLLAFLIILQKRVNORR 780  
Qy 781 MAQAFQNVREEPVAFNSGTTALNRKVKNPDPPTIYVVLNDNDIKFQDVIGEGNFGQVLK 840  
Db 781 MAQAFQNVREEPVAFNSGTTALNRKVKNPDPPTIYVVLNDNDIKFQDVIGEGNFGQVLK 840  
Qy 841 ARIKKGDLMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNINILGACEHRGYLYL 900  
Db 841 ARIKKGDLMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNINILGACEHRGYLYL 900  
Qy 901 ALEYAPHGNLLDFLRSVLETDPAFANSTASTLSQQLLHFAADVARGMDYLSQKQF 960  
Db 901 ALEYAPHGNLLDFLRSVLETDPAFANSTASTLSQQLLHFAADVARGMDYLSQKQF 960  
Qy 961 IHRDLAARNILVGENYVAKIADFGLSRQGVVYVKKTMGRLPVRMMAIESLNSYVYTTNSD 1020  
Db 961 IHRDLAARNILVGENYVAKIADFGLSRQGVVYVKKTMGRLPVRMMAIESLNSYVYTTNSD 1020  
Qy 1021 VMSYGVLLWEIVSLGTPYCGMTCAELYEKLPQGYRLEKPLNCDDEVDLMRQWREKPY 1080  
Db 1021 VMSYGVLLWEIVSLGTPYCGMTCAELYEKLPQGYRLEKPLNCDDEVDLMRQWREKPY 1080  
Qy 1081 ERPSFAQILVSLNRMLEBKTVNTLYEKFYVAGIDCSAEAA 1124  
Db 1081 ERPSFAQILVSLNRMLEBKTVNTLYEKFYVAGIDCSAEAA 1124

RESULT 4  
US-10-223-090-82  
; Sequence 82, Application US/10223090  
; Publication No. US20030105013A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Marsters, Scot A.  
; APPLICANT: Pan, James  
; APPLICANT: Stephan, Jean-Philippe P.  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Williams, P.Mickey  
; APPLICANT: Ye, Weilan  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; TREATMENT OF DISORDERS INVOLVING ANGIOGENESIS  
; FILE REFERENCE: P3235P1C2  
; CURRENT APPLICATION NUMBER: US/10/223,090  
; CURRENT FILING DATE: 2002-08-16  
; PRIOR APPLICATION NUMBER: US 10/081,056  
; PRIOR FILING DATE: 2002-02-20  
; PRIOR APPLICATION NUMBER: US 60/213,637  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: US 60/219,556  
; PRIOR FILING DATE: 2000-07-20  
; PRIOR APPLICATION NUMBER: US 60/220,624  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: US 60/220,664  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: PCT/US00/20710  
; PRIOR FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: US 60/222,695  
; PRIOR FILING DATE: 2000-08-02  
; PRIOR APPLICATION NUMBER: US 09/643,657  
; PRIOR FILING DATE: 2000-08-17  
; PRIOR APPLICATION NUMBER: PCT/US00/23522  
; PRIOR FILING DATE: 2000-08-23  
; PRIOR APPLICATION NUMBER: PCT/US00/23328  
; PRIOR FILING DATE: 2000-08-24  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 383  
; SEQ ID NO 82  
; LENGTH: 1124  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-223-090-82

Query Match 100.0%; Score 6013; DB 4; Length 1124;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MDSLASVLCVSLLSGTVGAMDLILINSLPLVSDAETSLTCTASGWRPHEPTIGRD 60  
Db 1 MDSLASVLCVSLLSGTVGAMDLILINSLPLVSDAETSLTCTASGWRPHEPTIGRD 60  
Qy 61 FEALMNQHDPLEVTQDVTREWAKVVMKREKASKINGAYFCEGRVGEAIRITMKMRQ 120  
Db 61 FEALMNQHDPLEVTQDVTREWAKVVMKREKASKINGAYFCEGRVGEAIRITMKMRQ 120  
Qy 121 QASFLPATLTMTVDKGNVNISFKKVLIKEBDATYKNGSFIHSVPRHEVPDILEVHLPH 180  
Db 121 QASFLPATLTMTVDKGNVNISFKKVLIKEBDATYKNGSFIHSVPRHEVPDILEVHLPH 180  
Qy 181 AQPQDAGVYSARYTGGNLTSAFTRLIVRRCEAQKWGECNHLCTACNNGVCHEDTGE 240  
Db 181 AQPQDAGVYSARYTGGNLTSAFTRLIVRRCEAQKWGECNHLCTACNNGVCHEDTGE 240  
Qy 241 ICPGFMGRTECEKACELHTFGRTCKERCSCGQCKSYVFCCLPDPYGCSCATGKGLQCNH 300  
Db 241 ICPGFMGRTECEKACELHTFGRTCKERCSCGQCKSYVFCCLPDPYGCSCATGKGLQCNH 300  
Qy 301 ACHPGFYGPDCKLRCSCNNGEMCDRFOGCLSPGQGLQCEBEGIPRMTPKIVDLPDHE 360  
Db 301 ACHPGFYGPDCKLRCSCNNGEMCDRFOGCLSPGQGLQCEBEGIPRMTPKIVDLPDHE 360  
Qy 361 VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAVFTTHRIILPPDSG 420  
Db 361 VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAVFTTHRIILPPDSG 420  
Qy 421 VVCSVNTVAGWVEKPFNISVKVLPRKPLNAPNVIDTGHNFVAVINTSSBYPFGDGPIKSKK 480

Db 421 VWCSTNTVAGWVERFNSISVVKPLNAPNVIDTGHFPAVINISSEBYFGDPIKSKK 480  
QY 481 LLYKPNHYEAWQHIOVTNEIIVTLNLETRTEYELCVQVRRGEGEGHPGVRPTTAS 540  
Db 481 LLYKPNHYEAWQHIOVTNEIIVTLNLETRTEYELCVQVRRGEGEGHPGVRPTTAS 540  
QY 541 IGLPPRGNLPLPKSOTTLNLTWQPIFPSSSEDDFYVEVERRSVQKSDQDNKVPGNLTSV 600  
Db 541 IGLPPRGNLPLPKSOTTLNLTWQPIFPSSSEDDFYVEVERRSVQKSDQDNKVPGNLTSV 600  
QY 601 LNNLHPRQYVVRARVNTKAQEWSEDTATWLSLILPPQENIKISNTHSSAVISWT 660  
Db 601 LNNLHPRQYVVRARVNTKAQEWSEDTATWLSLILPPQENIKISNTHSSAVISWT 660  
QY 661 ILDGYSISSITIRYKQGNEDQHDVVKIKNATIIQYQKLGLEPETAYQVDIFAENNIGS 720  
Db 661 ILDGYSISSITIRYKQGNEDQHDVVKIKNATIIQYQKLGLEPETAYQVDIFAENNIGS 720  
QY 721 SNPAFSHELVTLPESQAPADLGGKMLLTAIILGSAGMTCLTVLLAPLIILQLKRANVQRR 780  
Db 721 SNPAFSHELVTLPESQAPADLGGKMLLTAIILGSAGMTCLTVLLAPLIILQLKRANVQRR 780  
QY 781 MAQAFONVREEPAVQNSGTALNRKVKNPDPITYPVLDNDIKPQDVIGSGNFGVLK 840  
Db 781 MAQAFONVREEPAVQNSGTALNRKVKNPDPITYPVLDNDIKPQDVIGSGNFGVLK 840  
QY 841 ARIKODGLRMDAAIKEMKEYASKDHRDPAGELEVLCKLGHHPNIIINLLGACEHGRGYYL 900  
Db 841 ARIKODGLRMDAAIKEMKEYASKDHRDPAGELEVLCKLGHHPNIIINLLGACEHGRGYYL 900  
QY 901 AIEYAPHGNLDFLRKSRVLETPDPAFIANSTASTLSSQOLLHFAADVARGM DYLSQKQF 960  
Db 901 AIEYAPHGNLDFLRKSRVLETPDPAFIANSTASTLSSQOLLHFAADVARGM DYLSQKQF 960  
QY 961 IHRDLAARNILGVENYAKIADFLSRGQVYVYKTMGBLPVRWMAIESLNTSVYTTNSD 1020  
Db 961 IHRDLAARNILGVENYAKIADFLSRGQVYVYKTMGBLPVRWMAIESLNTSVYTTNSD 1020  
QY 1021 VMSGYVLLWEIVSLGTPYCGMTCAELYEKLPQGYRLEKPLNCDDBVYDLMRQCWREKPY 1080  
Db 1021 VMSGYVLLWEIVSLGTPYCGMTCAELYEKLPQGYRLEKPLNCDDBVYDLMRQCWREKPY 1080  
QY 1081 ERPSFAQIILVSLNRMLEERKTYVNTTLYEKFTYAGIDCSAEAAA 1124  
Db 1081 ERPSFAQIILVSLNRMLEERKTYVNTTLYEKFTYAGIDCSAEAAA 1124

RESULT 5

US-10-223-087-82  
; Sequence 82, Application US/10223087  
; Publication No. US20030109438A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Marsters, Scott A.  
; APPLICANT: Pan, James  
; APPLICANT: Stephan, Jean-Philippe F.  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Ye, Weilan  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; FILE REFERENCE: P3235PIC4  
; CURRENT APPLICATION NUMBER: US/10/223,087  
; CURRENT FILING DATE: 2002-08-16

; PRIOR APPLICATION NUMBER: US 10/081,056  
; PRIOR FILING DATE: 2002-02-20  
; PRIOR APPLICATION NUMBER: US 60/213,637  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: US 60/219,556  
; PRIOR FILING DATE: 2000-07-20  
; PRIOR APPLICATION NUMBER: US 60/220,624  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: US 60/220,664  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: PCT/US00/20710  
; PRIOR FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: US 60/222,695  
; PRIOR FILING DATE: 2000-08-02  
; PRIOR APPLICATION NUMBER: US 09/643,657  
; PRIOR FILING DATE: 2000-08-17  
; PRIOR APPLICATION NUMBER: PCT/US00/23522  
; PRIOR FILING DATE: 2000-08-23  
; PRIOR APPLICATION NUMBER: PCT/US00/23328  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/230,978  
; PRIOR FILING DATE: 2000-09-07  
; PRIOR APPLICATION NUMBER: US 60/232,887  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 09/664,610  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: US 09/665,350  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: US 60/242,922  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 09/709,238  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: PCT/US00/30952  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: PCT/US00/30873  
; PRIOR FILING DATE: 2000-11-10  
; PRIOR APPLICATION NUMBER: PCT/US00/32678  
; PRIOR FILING DATE: 2000-12-01  
; PRIOR APPLICATION NUMBER: US 09/747,259  
; PRIOR FILING DATE: 2000-12-20  
; PRIOR APPLICATION NUMBER: PCT/US00/34956  
; PRIOR FILING DATE: 2000-12-20  
; PRIOR APPLICATION NUMBER: US 09/767,609  
; PRIOR FILING DATE: 2001-01-22  
; PRIOR APPLICATION NUMBER: US 09/796,498  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: PCT/US01/06520  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: PCT/US01/06666  
; PRIOR FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: US 09/802,706  
; PRIOR FILING DATE: 2001-03-09  
; PRIOR APPLICATION NUMBER: US 09/808,689  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: US 09/816,744  
; PRIOR FILING DATE: 2001-03-22  
; PRIOR APPLICATION NUMBER: US 09/828,366  
; PRIOR FILING DATE: 2001-04-05  
; PRIOR APPLICATION NUMBER: US 09/854,208  
; PRIOR FILING DATE: 2001-05-10  
; PRIOR APPLICATION NUMBER: US 09/854,280  
; PRIOR FILING DATE: 2001-05-10  
; PRIOR APPLICATION NUMBER: US 09/866,028  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: US 09/866,034  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: PCT/US01/17092  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: US 09/870,574  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: PCT/US01/17443  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: PCT/US01/17800

```

; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US01/19692
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: PCT/US01/21066
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: PCT/US01/21735
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 383
; SEQ ID NO 82
; LENGTH: 1124
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-223-087-82

Query Match      100.0%; Score 6013; DB 4; Length 1124;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLASLVLGVSLLSGTVEGAMDILILNSLPLVSDAFTSLTCIASGWRPHEPITIGRD 60
DB 1 MDLASLVLGVSLLSGTVEGAMDILILNSLPLVSDAFTSLTCIASGWRPHEPITIGRD 60

QY 61 FEALMNQHDPLEVTDVTVREWAKVWVKREKASKINGAYFCEGRVGRGEAIRITMKMRQ 120
DB 61 FEALMNQHDPLEVTDVTVREWAKVWVKREKASKINGAYFCEGRVGRGEAIRITMKMRQ 120

QY 121 QASFLPATLTMVTVDKGDVNIISFKVLIKEEDAVLYKXGSPFIHSVPRHEVPDILEVHLPH 180
DB 121 QASFLPATLTMVTVDKGDVNIISFKVLIKEEDAVLYKXGSPFIHSVPRHEVPDILEVHLPH 180

QY 181 AQPDAGVYSARYIIGNLFTSAFTRLIVRCEAQKWGPCNHLCTACMNGVYCHEDTGC 240
DB 181 AQPDAGVYSARYIIGNLFTSAFTRLIVRCEAQKWGPCNHLCTACMNGVYCHEDTGC 240

QY 241 ICPPGFMGRTCKACELHTFGRTCKERCSCGQCKSYVFCPLDPPYGCSCATGKGLQNE 300
DB 241 ICPPGFMGRTCKACELHTFGRTCKERCSCGQCKSYVFCPLDPPYGCSCATGKGLQNE 300

QY 301 ACHPGFYGDCKLRCSNNGEMCDRFGQCLCPGQGLQCEREGIPRMTPKIVDLPDHE 360
DB 301 ACHPGFYGDCKLRCSNNGEMCDRFGQCLCPGQGLQCEREGIPRMTPKIVDLPDHE 360

QY 361 VNSGKFNPTCKASGWLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAFTTHRIILPPDSG 420
DB 361 VNSGKFNPTCKASGWLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAFTTHRIILPPDSG 420

QY 421 VVWCSVNTVAGWVEKPFNISKVLKPLNAPNVIDTGHNFVNINISSEBPFYFGDGPIKSK 480
DB 421 VVWCSVNTVAGWVEKPFNISKVLKPLNAPNVIDTGHNFVNINISSEBPFYFGDGPIKSK 480

QY 481 LLYKPVNHYEAWQHIQVNEIVTLNVLPRTEYELCVOLVRGEGEHGPVREFTTAS 540
DB 481 LLYKPVNHYEAWQHIQVNEIVTLNVLPRTEYELCVOLVRGEGEHGPVREFTTAS 540

QY 541 IGLPPRGNLNLPKSTQTLNLTWQIFPSSSEDDFYVEVERRSVQKSDQONIKVPGNLTSV 600
DB 541 IGLPPRGNLNLPKSTQTLNLTWQIFPSSSEDDFYVEVERRSVQKSDQONIKVPGNLTSV 600

QY 601 LLNLNHPREYVVRARVNTKAGEWSEDLTAWTSLDILPPQENIKISNITHSSAVISWT 660
DB 601 LLNLNHPREYVVRARVNTKAGEWSEDLTAWTSLDILPPQENIKISNITHSSAVISWT 660

QY 661 ILDGYSISITIRYKVGKNEQHDVVKIKNATIIQYOLKGLEPETAYQVDIFPAENNIGS 720
DB 661 ILDGYSISITIRYKVGKNEQHDVVKIKNATIIQYOLKGLEPETAYQVDIFPAENNIGS 720

QY 721 SNPAFSELVTLPEAQADLGGGKMLJIALIGSAGMTCLTVLLAFLIILQKRVNQR 780
DB 721 SNPAFSELVTLPEAQADLGGGKMLJIALIGSAGMTCLTVLLAFLIILQKRVNQR 780

QY 781 MAQAFQNVREEPVAFQNSGTLALNRKVNPPDTIYPVLDWMDIKFQDVIGEGNFGVLK 840
DB 781 MAQAFQNVREEPVAFQNSGTLALNRKVNPPDTIYPVLDWMDIKFQDVIGEGNFGVLK 840
```

```

QY 841 ARIKKGRLMDAAIKRMKEYASKDHRDFAGELEVLCKLGHHPNIIINLIGACEHRGYLYL 900
DB 841 ARIKKGRLMDAAIKRMKEYASKDHRDFAGELEVLCKLGHHPNIIINLIGACEHRGYLYL 900

QY 901 AIEYAPHGNLLDFLRKSRVLETPDPAFAIANSTASTLSSQQLLHFAADVARGMDYLSQKF 960
DB 901 AIEYAPHGNLLDFLRKSRVLETPDPAFAIANSTASTLSSQQLLHFAADVARGMDYLSQKF 960

QY 961 IHRDLAARNIILVGENYVAKIADFGLSRQGEVYVVKTMGRLPVRWMAIESLNVSVYTTNSD 1020
DB 961 IHRDLAARNIILVGENYVAKIADFGLSRQGEVYVVKTMGRLPVRWMAIESLNVSVYTTNSD 1020

QY 1021 VMSYGVLLWEIVSLGCTPYCGMTCAELYEKLPQGYRLEKPLNCDDDEVYDLMRQWREKPY 1080
DB 1021 VMSYGVLLWEIVSLGCTPYCGMTCAELYEKLPQGYRLEKPLNCDDDEVYDLMRQWREKPY 1080

QY 1081 ERPSFAQILVSLNRMLEERKTYVNTTLYEKFTYAGIDCSAEAA 1124
DB 1081 ERPSFAQILVSLNRMLEERKTYVNTTLYEKFTYAGIDCSAEAA 1124
```

RESULT 6

```

US-10-223-083-82
; Sequence 82, Application US/10223083
; Publication No. US20030119112A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Marsters, Scot A.
; APPLICANT: Pan, James
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Ye, Weilan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: P3235P1C8
; CURRENT APPLICATION NUMBER: US/10/223,083
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 10/081,056
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: US 60/213,637
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/219,556
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: US 60/220,624
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/220,664
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: PCT/US00/20710
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/222,695
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: US 09/643,657
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: PCT/US00/23522
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 383
; SEQ ID NO 82
; LENGTH: 1124
; TYPE: PRT
; ORGANISM: Homo sapiens
```

US-10-223-083-82																																																						
Query Match 100.0%; Score 6013; DB 4; Length 1124;																																																						
Best Local Similarity 100.0%; Pred. No. 0;																																																						
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;																																																						
QY	1	MDSLASLVL	CGVSLLS	GTVEG	AMDLIL	INSPLV	SDAETS	LTCTAS	GWRRPHEPITIGRD 60																																													
DB	1	MDSLASLVL	CGVSLLS	GTVEG	AMDLIL	INSPLV	SDAETS	LTCTAS	GWRRPHEPITIGRD 60																																													
QY	61	FEALMNQ	ODPLEV	TQDVT	REWAKV	WVKREK	AS	INGAYF	CEGVRVGEAIRIRTMKMQ 120																																													
DB	61	FEALMNQ	ODPLEV	TQDVT	REWAKV	WVKREK	AS	INGAYF	CEGVRVGEAIRIRTMKMQ 120																																													
QY	121	QASFLPAT	LTMVTD	KGNVNI	ISFKVL	IKKEEDA	VIYKNGS	FIHSVPRHEVPDILEVHLPH 180																																														
DB	121	QASFLPAT	LTMVTD	KGNVNI	ISFKVL	IKKEEDA	VIYKNGS	FIHSVPRHEVPDILEVHLPH 180																																														
QY	181	AQPDAGVY	SARYIG	NLFTSA	FTRLIV	VRCEAQ	KWGPEC	NHLC	TACMNGVCHEDTGC 240																																													
DB	181	AQPDAGVY	SARYIG	NLFTSA	FTRLIV	VRCEAQ	KWGPEC	NHLC	TACMNGVCHEDTGC 240																																													
QY	241	ICPPGFMGT	CEKACEL	HFTG	TKERC	SGQGS	YVFC	LPDP	PGSCATGKWLQNE 300																																													
DB	241	ICPPGFMGT	CEKACEL	HFTG	TKERC	SGQGS	YVFC	LPDP	PGSCATGKWLQNE 300																																													
QY	301	ACHPGFYG	PDCKLR	CSNNGE	CMDFQ	GC	SPGQ	GLC	REGIPRMTPKIVDLPDHIE 360																																													
DB	301	ACHPGFYG	PDCKLR	CSNNGE	CMDFQ	GC	SPGQ	GLC	REGIPRMTPKIVDLPDHIE 360																																													
QY	361	VNSGKFN	PKICKAS	GWPL	PTNEB	MTLV	KPDG	TVLHPK	DNHTDHFVAITFTHRIILPPDSG 420																																													
DB	361	VNSGKFN	PKICKAS	GWPL	PTNEB	MTLV	KPDG	TVLHPK	DNHTDHFVAITFTHRIILPPDSG 420																																													
QY	421	VWCVSVNT	VAGVVE	PFNLSV	KPLNAP	NVID	TGHN	FAVIN	SSBYPFGDGPIKSKK 480																																													
DB	421	VWCVSVNT	VAGVVE	PFNLSV	KPLNAP	NVID	TGHN	FAVIN	SSBYPFGDGPIKSKK 480																																													
QY	481	LLYKPNHYE	AWQHI	QVTNEI	VTLNLE	PRTEYEL	CVQ	LVRRGE	GEHGPVRRFTAS 540																																													
DB	481	LLYKPNHYE	AWQHI	QVTNEI	VTLNLE	PRTEYEL	CVQ	LVRRGE	GEHGPVRRFTAS 540																																													
QY	541	IGLPPRG	LNLLPK	SQTTLN	LWQPI	FPSS	EDDF	YVE	VERRSVQSDQONIKVPGNLSV 600																																													
DB	541	IGLPPRG	LNLLPK	SQTTLN	LWQPI	FPSS	EDDF	YVE	VERRSVQSDQONIKVPGNLSV 600																																													
QY	601	LNUNLHP	REYQYV	RVARVNT	KAQWSE	DLTAW	LS	DILP	POPENIKISNI	THSSAVISWT 660																																												
DB	601	LNUNLHP	REYQYV	RVARVNT	KAQWSE	DLTAW	LS	DILP	POPENIKISNI	THSSAVISWT 660																																												
QY	661	IUDGYS	ISITIRYK	VQGNED	QHV	DKI	KNATI	IYQ	LKLE	PETAYQVDIPAE	NNIGS 720																																											
DB	661	IUDGYS	ISITIRYK	VQGNED	QHV	DKI	KNATI	IYQ	LKLE	PETAYQVDIPAE	NNIGS 720																																											
QY	721	SNPAF	SHELVTLP	ESQAP	ADLGG	KMLL	IA	ILGS	AGMTCL	TVLLAFL	IIQLK	RANVQRR 780																																										
DB	721	SNPAF	SHELVTLP	ESQAP	ADLGG	KMLL	IA	ILGS	AGMTCL	TVLLAFL	IIQLK	RANVQRR 780																																										
QY	781	MAQAF	QNVRE	BAVQ	PN	SGT	LALNR	KVN	PNPDT	YV	LDND	IKFQ	ODVIG	EGNFQVULK 840																																								
DB	781	MAQAF	QNVRE	BAVQ	PN	SGT	LALNR	KVN	PNPDT	YV	LDND	IKFQ	ODVIG	EGNFQVULK 840																																								
QY	841	ARIK	KG	LRMD	AAIK	KMK	EYAS	KDD	HRD	FAGE	LV	CKL	GHP	NI	INLL	GAC	EHGYLYL 900																																					
DB	841	ARIK	KG	LRMD	AAIK	KMK	EYAS	KDD	HRD	FAGE	LV	CKL	GHP	NI	INLL	GAC	EHGYLYL 900																																					
QY	901	AIEY	A	PHGN	LLD	FL	RS	R	SVLE	TD	P	A	I	A	N	S	T	A	S	T	L	S	S	Q	L	L	H	F	A	A	D	V	A	R	G	M	D	Y	L	S	Q	K	F 960											
DB	901	AIEY	A	PHGN	LLD	FL	RS	R	SVLE	TD	P	A	I	A	N	S	T	A	S	T	L	S	S	Q	L	L	H	F	A	A	D	V	A	R	G	M	D	Y	L	S	Q	K	F 960											
QY	961	IHRD	L	A	A	R	N	I	L	G	E	N	T	V	A	K	I	A	D	F	G	L	S	R	G	O	E	V	V	Y	K	K	T	M	G	R	L	P	V	R	W	A	I	E	S	L	N	S	V	T	T	N	S	D 1020
DB	961	IHRD	L	A	A	R	N	I	L	G	E	N	T	V	A	K	I	A	D	F	G	L	S	R	G	O	E	V	V	Y	K	K	T	M	G	R	L	P	V	R	W	A	I	E	S	L	N	S	V	T	T	N	S	D 1020

QY	1021	VNSGVLLWEI	VS	LG	TPY	CG	MT	CAE	LYE	KL	PO	GY	R	L	E	K	P	NC	D	D	E	V	Y	D	L	M	R	O	C	W	R	E	K	P	Y 1080		
DB	1021	VNSGVLLWEI	VS	LG	TPY	CG	MT	CAE	LYE	KL	PO	GY	R	L	E	K	P	NC	D	D	E	V	Y	D	L	M	R	O	C	W	R	E	K	P	Y 1080		
QY	1081	ERP	S	PA	Q	L	V	S	L	N	R	M	L	B	E	R	K	T	Y	N	T	T	L	E	K	F	T	A	G	I	D	C	S	A	E	E	A 1124
DB	1081	ERP	S	PA	Q	L	V	S	L	N	R	M	L	B	E	R	K	T	Y	N	T	T	L	E	K	F	T	A	G	I	D	C	S	A	E	E	A 1124

RESULT 7

US-10-223-089-82

Sequence 82, Application US/10223089

Publication No. US20030125521A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Ferrara, Napoleone

APPLICANT: Gerber, Hanspeter

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Hillan, Kenneth J.

APPLICANT: Masters, Scot A.

APPLICANT: Pan, James

APPLICANT: Stephan, Jean-Philippe F.

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.

APPLICANT: Williams, P.Mickey

APPLICANT: Ye, Weilan

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF DISORDERS INVOLVING ANGIOGENESIS

FILE REFERENCE: P3235P1C9

CURRENT APPLICATION NUMBER: US/10/223,089

CURRENT FILING DATE: 2002-08-16

PRIOR APPLICATION NUMBER: US 10/081,056

PRIOR FILING DATE: 2002-02-20

PRIOR APPLICATION NUMBER: US 60/213,637

PRIOR FILING DATE: 2000-06-23

PRIOR APPLICATION NUMBER: US 60/219,556

PRIOR FILING DATE: 2000-07-20

PRIOR APPLICATION NUMBER: US 60/220,624

PRIOR FILING DATE: 2000-07-25

PRIOR APPLICATION NUMBER: US 60/220,664

PRIOR FILING DATE: 2000-07-25

PRIOR APPLICATION NUMBER: PCT/US00/20710

PRIOR FILING DATE: 2000-07-28

PRIOR APPLICATION NUMBER: US 60/222,695

PRIOR FILING DATE: 2000-08-02

PRIOR APPLICATION NUMBER: US 09/643,657

PRIOR APPLICATION NUMBER: PCT/US00/23522

PRIOR FILING DATE: 2000-08-23

PRIOR APPLICATION NUMBER: PCT/US00/23328

PRIOR FILING DATE: 2000-08-24

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 383

SEQ ID NO 82

LENGTH: 1124

TYPE: PRT

ORGANISM: Homo sapiens

US-10-223-089-82

Query Match 100.0%; Score 6013; DB 4; Length 1124;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MDSLASLVL	CGVSLLS	GTVEG	AMDLIL	INSPLV	SDAETS	LTCTAS	GWRRPHEPITIGRD 60																																													
DB	1	MDSLASLVL	CGVSLLS	GTVEG	AMDLIL	INSPLV	SDAETS	LTCTAS	GWRRPHEPITIGRD 60																																													
QY	61	FEALMNQ	ODPLEV	TQDVT	REWAKV	WVKREK	AS	INGAYF	CEGVRVGEAIRIRTMKMQ 120																																													
DB	61	FEALMNQ	ODPLEV	TQDVT	REWAKV	WVKREK	AS	INGAYF	CEGVRVGEAIRIRTMKMQ 120																																													
QY	121	QASFLPAT	LTMVTD	KGNVNI	ISFKVL	IKKEEDA	VIYKNGS	FIHSVPRHEVPDILEVHLPH 180																																														
DB	121	QASFLPAT	LTMVTD	KGNVNI	ISFKVL	IKKEEDA	VIYKNGS	FIHSVPRHEVPDILEVHLPH 180																																														
QY	181	AQPDAGVY	SARYIG	NLFTSA	FTRLIV	VRCEAQ	KWGPEC	NHLC	TACMNGVCHEDTGC 240																																													
DB	181	AQPDAGVY	SARYIG	NLFTSA	FTRLIV	VRCEAQ	KWGPEC	NHLC	TACMNGVCHEDTGC 240																																													
QY	241	ICPPGFMGT	CEKACEL	HFTG	TKERC	SGQGS	YVFC	LPDP	PGSCATGKWLQNE 300																																													
DB	241	ICPPGFMGT	CEKACEL	HFTG	TKERC	SGQGS	YVFC	LPDP	PGSCATGKWLQNE 300																																													
QY	301	ACHPGFYG	PDCKLR	CSNNGE	CMDFQ	GC	SPGQ	GLC	REGIPRMTPKIVDLPDHIE 360																																													
DB	301	ACHPGFYG	PDCKLR	CSNNGE	CMDFQ	GC	SPGQ	GLC	REGIPRMTPKIVDLPDHIE 360																																													
QY	361	VNSGKFN	PKICKAS	GWPL	PTNEB	MTLV	KPDG	TVLHPK	DNHTDHFVAITFTHRIILPPDSG 420																																													
DB	361	VNSGKFN	PKICKAS	GWPL	PTNEB	MTLV	KPDG	TVLHPK	DNHTDHFVAITFTHRIILPPDSG 420																																													
QY	421	VWCVSVNT	VAGVVE	PFNLSV	KPLNAP	NVID	TGHN	FAVIN	SSBYPFGDGPIKSKK 480																																													
DB	421	VWCVSVNT	VAGVVE	PFNLSV	KPLNAP	NVID	TGHN	FAVIN	SSBYPFGDGPIKSKK 480																																													
QY	481	LLYKPNHYE	AWQHI	QVTNEI	VTLNLE	PRTEYEL	CVQ	LVRRGE	GEHGPVRRFTAS 540																																													
DB	481	LLYKPNHYE	AWQHI	QVTNEI	VTLNLE	PRTEYEL	CVQ	LVRRGE	GEHGPVRRFTAS 540																																													
QY	541	IGLPPRG	LNLLPK	SQTTLN	LWQPI	FPSS	EDDF	YVE	VERRSVQSDQONIKVPGNLSV 600																																													
DB	541	IGLPPRG	LNLLPK	SQTTLN	LWQPI	FPSS	EDDF	YVE	VERRSVQSDQONIKVPGNLSV 600																																													
QY	601	LNUNLHP	REYQYV	RVARVNT	KAQWSE	DLTAW	LS	DILP	POPENIKISNI	THSSAVISWT 660																																												
DB	601	LNUNLHP	REYQYV	RVARVNT	KAQWSE	DLTAW	LS	DILP	POPENIKISNI	THSSAVISWT 660																																												
QY	661	IUDGYS	ISITIRYK	VQGNED	QHV	DKI	KNATI	IYQ	LKLE	PETAYQVDIPAE	NNIGS 720																																											
DB	661	IUDGYS	ISITIRYK	VQGNED	QHV	DKI	KNATI	IYQ	LKLE	PETAYQVDIPAE	NNIGS 720																																											
QY	721	SNPAF	SHELVTLP	ESQAP	ADLGG	KMLL	IA	ILGS	AGMTCL	TVLLAFL	IIQLK	RANVQRR 780																																										
DB	721	SNPAF	SHELVTLP	ESQAP	ADLGG	KMLL	IA	ILGS	AGMTCL	TVLLAFL	IIQLK	RANVQRR 780																																										
QY	781	MAQAF	QNVRE	BAVQ	PN	SGT	LALNR	KVN	PNPDT	YV	LDND	IKFQ	ODVIG	EGNFQVULK 840																																								
DB	781	MAQAF	QNVRE	BAVQ	PN	SGT	LALNR	KVN	PNPDT	YV	LDND	IKFQ	ODVIG	EGNFQVULK 840																																								
QY	841	ARIK	KG	LRMD	AAIK	KMK	EYAS	KDD	HRD	FAGE	LV	CKL	GHP	NI	INLL	GAC	EHGYLYL 900																																					
DB	841	ARIK	KG	LRMD	AAIK	KMK	EYAS	KDD	HRD	FAGE	LV	CKL	GHP	NI	INLL	GAC	EHGYLYL 900																																					
QY	901	AIEY	A	PHGN	LLD	FL	RS	R	SVLE	TD	P	A	I	A	N	S	T	A	S	T	L	S	S	Q	L	L	H	F	A	A	D	V	A	R	G	M	D	Y	L	S	Q	K	F 960											
DB	901	AIEY	A	PHGN	LLD	FL	RS	R	SVLE	TD	P	A	I	A	N	S	T	A	S	T	L	S	S	Q	L	L	H	F	A	A	D	V	A	R	G	M	D	Y	L	S	Q	K	F 960											
QY	961	IHRD	L	A	A	R	N	I	L	G	E	N	T	V	A	K	I	A	D	F	G	L	S	R	G	O	E	V	V	Y	K	K	T	M	G	R	L	P	V	R	W	A	I	E	S	L	N	S	V	T	T	N	S	D 1020
DB	961	IHRD	L	A	A	R	N	I	L	G	E	N	T	V	A	K	I	A	D	F	G	L	S	R	G	O	E	V	V	Y	K	K	T	M	G	R	L	P	V	R	W	A	I	E	S	L	N	S	V	T	T	N	S	D 1020

US-10-223-089-82									
; Sequence 82, Application US/10223089									
; Publication No. US2003012521A1									
; GENERAL INFORMATION:									
; APPLICANT: Baker, Kevin P.									
; APPLICANT: Ferrara, Napoleone									
; APPLICANT: Gerber, Hanspeter									
; APPLICANT: Gerritsen, Mary E.									
; APPLICANT: Goddard, Audrey									
; APPLICANT: Godowski, Paul J.									
; APPLICANT: Gurney, Austin L.									
; APPLICANT: Hillan, Kenneth J.									
; APPLICANT: Marsters, Scot A.									
; APPLICANT: Pan, James									
; APPLICANT: Stephan, Jean-Philippe F.									
; APPLICANT: Watanabe, Colin K.									
; APPLICANT: Wood, William I.									
; APPLICANT: Williams, P.Mickey									
; APPLICANT: Ye, Weilan									
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND									
; TITLE OF INVENTION: TREATMENT OF DISORDERS INVOLVING ANGIOGENESIS									
; FILE REFERENCE: P3235P1C9									
; CURRENT APPLICATION NUMBER: US/10/223,089									
; CURRENT FILING DATE: 2002-08-15									
; PRIOR APPLICATION NUMBER: US 10/081,056									
; PRIOR FILING DATE: 2002-02-20									
; PRIOR APPLICATION NUMBER: US 60/213,637									
; PRIOR FILING DATE: 2000-06-23									
; PRIOR APPLICATION NUMBER: US 60/219,556									
; PRIOR FILING DATE: 2000-07-20									
; PRIOR APPLICATION NUMBER: US 60/220,624									
; PRIOR FILING DATE: 2000-07-25									
; PRIOR APPLICATION NUMBER: US 60/220,664									
; PRIOR FILING DATE: 2000-07-25									
; PRIOR APPLICATION NUMBER: PCT/US00/20710									
; PRIOR FILING DATE: 2000-07-28									
; PRIOR APPLICATION NUMBER: US 60/222,695									
; PRIOR FILING DATE: 2000-08-02									
; PRIOR APPLICATION NUMBER: US 09/643,657									
; PRIOR FILING DATE: 2000-08-17									
; PRIOR APPLICATION NUMBER: PCT/US00/23522									
; PRIOR FILING DATE: 2000-08-23									
; PRIOR APPLICATION NUMBER: PCT/US00/23328									
; PRIOR FILING DATE: 2000-08-24									
; Remaining Prior Application data removed - See File Wrapper or PALM.									
; NUMBER OF SEQ ID NOS: 383									
; SEQ ID NO 82									
; LENGTH: 1124									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-10-223-089-82									
Query Match 100.0%; Score 6013; DB 4; Length 1124;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MDSLASLVL	CGVSLLS	SGTVEG	AMDLIL	INSPLV	SDAETS	LTCTAS	GWRRPHEPITIGRD 60
DB	1	MDSLASLVL	CGVSLLS	SGTVEG	AMDLIL	INSPLV	SDAETS	LTCTAS	GWRRPHEPITIGRD 60
QY	61	FEALMNQ	ODPLEV	TQDVT	REWAKV	WVKREK	AS	INGAYF	CEGVRVGEAIRIRTMKMQ 120

```
Db 61 FEALMNQHDPLEVTQDVTREWAKKVVWKREKASKINGAYFCEGRVVRGEAIRRTMKMRQ 120
Qy 121 QASFLPATLTMTVDKGNVNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILVHLPH 180
Db 121 QASFLPATLTMTVDKGNVNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILVHLPH 180
Qy 181 AQPQDAGVYSARYIGNLTFTSAFTRLIVRRCEAKWGPCNHLCTACWNGVCHEDTGC 240
Db 181 AQPQDAGVYSARYIGNLTFTSAFTRLIVRRCEAKWGPCNHLCTACWNGVCHEDTGC 240
Qy 241 ICPPGFMGRTCEKACELHTFGRTCKERCSGQGCCKSVYVFCCLPDPYGCSCATGWKGLQNE 300
Db 241 ICPPGFMGRTCEKACELHTFGRTCKERCSGQGCCKSVYVFCCLPDPYGCSCATGWKGLQNE 300
Qy 301 ACHPGFYGPDCKLRCSNNGEMCDRFOGCLCSPGWQGLQCEREGIPRMTPKIVLDPDHE 360
Db 301 ACHPGFYGPDCKLRCSNNGEMCDRFOGCLCSPGWQGLQCEREGIPRMTPKIVLDPDHE 360
Qy 361 VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAVFTIHRILPPDSG 420
Db 361 VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAVFTIHRILPPDSG 420
Qy 421 VVWCSVNTVAGWVEKPFNISVKVLPKPLNAPNVIDTGHNFVAVINISSEBPYFGDGPIKSKK 480
Db 421 VVWCSVNTVAGWVEKPFNISVKVLPKPLNAPNVIDTGHNFVAVINISSEBPYFGDGPIKSKK 480
Qy 481 LLYKPVNHYEAWOHIQVNTNEIVTLNLYPRTEYELCVQLVRRGEGBGHPGVRFTTAS 540
Db 481 LLYKPVNHYEAWOHIQVNTNEIVTLNLYPRTEYELCVQLVRRGEGBGHPGVRFTTAS 540
Qy 541 IGLPPRGNLNLPKSOQTLNLTWQPIFPSSDDDFVEVERRSVQKSDQONIKVPGNLTSV 600
Db 541 IGLPPRGNLNLPKSOQTLNLTWQPIFPSSDDDFVEVERRSVQKSDQONIKVPGNLTSV 600
Qy 601 LNNLHPREQYVVRVAVNTKAGWSEDLTAWTSLDILPPOPENIKISNIHTSSAVISWT 660
Db 601 LNNLHPREQYVVRVAVNTKAGWSEDLTAWTSLDILPPOPENIKISNIHTSSAVISWT 660
Qy 661 ILDGSISITIRYKQGNEDQHVDKIKNATIIQYQLKGLEPETAYQVDIFAENNIGS 720
Db 661 ILDGSISITIRYKQGNEDQHVDKIKNATIIQYQLKGLEPETAYQVDIFAENNIGS 720
Qy 721 SNPAFSELVTLPEAOAPADLGGGKMLLIIAGSAGMTCLTVLLAFLIILQKRVANVORR 780
Db 721 SNPAFSELVTLPEAOAPADLGGGKMLLIIAGSAGMTCLTVLLAFLIILQKRVANVORR 780
Qy 781 MAQAFQNVREEPAVQFNSGTLALNRKVNPNPDPTIYPVLDMNDIKFQDVIGEGNGQVLK 840
Db 781 MAQAFQNVREEPAVQFNSGTLALNRKVNPNPDPTIYPVLDMNDIKFQDVIGEGNGQVLK 840
Qy 841 ARIKQGLRMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNTINLLGACEHRGYLYL 900
Db 841 ARIKQGLRMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNTINLLGACEHRGYLYL 900
Qy 901 ALEYAPHGNLLDPLKRSVLETPDPAFANSTASTLSOQLLHFAADVARGMDYLSQKF 960
Db 901 ALEYAPHGNLLDPLKRSVLETPDPAFANSTASTLSOQLLHFAADVARGMDYLSQKF 960
Qy 961 IHRDLAARNILVGENYAKIADFGLSRGQEVVYKKTGMRLPVRWMAIESLNSVYVTTNSD 1020
Db 961 IHRDLAARNILVGENYAKIADFGLSRGQEVVYKKTGMRLPVRWMAIESLNSVYVTTNSD 1020
Qy 1021 VWSYGVLLWEIVSLGTPYCGMTCAELYEKLPQGRLEKPLNCDDEVDLMRQCHREKPY 1080
Db 1021 VWSYGVLLWEIVSLGTPYCGMTCAELYEKLPQGRLEKPLNCDDEVDLMRQCHREKPY 1080
Qy 1081 ERPSFAQILVSNRMLEERKTVNTTLYEKFTYAGIDCSAEBA 1124
Db 1081 ERPSFAQILVSNRMLEERKTVNTTLYEKFTYAGIDCSAEBA 1124
```

```
; Sequence 112, Application US/10021660
; Publication No. US20030152926A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: No. US20030152926A1el Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021,660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 112
; LENGTH: 1124
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-021-660-112
```

Query Match 100.0%; Score 6013; DB 4; Length 1124;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 MDSLASVLCVSGVLLSGTVEGAMDLLILINSLPLVSDAETSLTCIASGWRPHEPITIGRD 60
Db 1 MDSLASVLCVSGVLLSGTVEGAMDLLILINSLPLVSDAETSLTCIASGWRPHEPITIGRD 60
Qy 61 FEALMNQHDPLEVTQDVTREWAKKVVWKREKASKINGAYFCEGRVVRGEAIRRTMKMRQ 120
Db 61 FEALMNQHDPLEVTQDVTREWAKKVVWKREKASKINGAYFCEGRVVRGEAIRRTMKMRQ 120
Qy 121 QASFLPATLTMTVDKGNVNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILVHLPH 180
Db 121 QASFLPATLTMTVDKGNVNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILVHLPH 180
Qy 181 AQPQDAGVYSARYIGNLTFTSAFTRLIVRRCEAKWGPCNHLCTACWNGVCHEDTGC 240
Db 181 AQPQDAGVYSARYIGNLTFTSAFTRLIVRRCEAKWGPCNHLCTACWNGVCHEDTGC 240
Qy 241 ICPPGFMGRTCEKACELHTFGRTCKERCSGQGCCKSVYVFCCLPDPYGCSCATGWKGLQNE 300
Db 241 ICPPGFMGRTCEKACELHTFGRTCKERCSGQGCCKSVYVFCCLPDPYGCSCATGWKGLQNE 300
Qy 301 ACHPGFYGPDCKLRCSNNGEMCDRFOGCLCSPGWQGLQCEREGIPRMTPKIVLDPDHE 360
Db 301 ACHPGFYGPDCKLRCSNNGEMCDRFOGCLCSPGWQGLQCEREGIPRMTPKIVLDPDHE 360
Qy 361 VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAVFTIHRILPPDSG 420
Db 361 VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAVFTIHRILPPDSG 420
Qy 421 VVWCSVNTVAGWVEKPFNISVKVLPKPLNAPNVIDTGHNFVAVINISSEBPYFGDGPIKSKK 480
Db 421 VVWCSVNTVAGWVEKPFNISVKVLPKPLNAPNVIDTGHNFVAVINISSEBPYFGDGPIKSKK 480
Qy 481 LLYKPVNHYEAWOHIQVNTNEIVTLNLYPRTEYELCVQLVRRGEGBGHPGVRFTTAS 540
Db 481 LLYKPVNHYEAWOHIQVNTNEIVTLNLYPRTEYELCVQLVRRGEGBGHPGVRFTTAS 540
Qy 541 IGLPPRGNLNLPKSOQTLNLTWQPIFPSSDDDFVEVERRSVQKSDQONIKVPGNLTSV 600
Db 541 IGLPPRGNLNLPKSOQTLNLTWQPIFPSSDDDFVEVERRSVQKSDQONIKVPGNLTSV 600
Qy 601 LNNLHPREQYVVRVAVNTKAGWSEDLTAWTSLDILPPOPENIKISNIHTSSAVISWT 660
Db 601 LNNLHPREQYVVRVAVNTKAGWSEDLTAWTSLDILPPOPENIKISNIHTSSAVISWT 660
```

QY 661 ILDGYSISSITIRYKVGKNEHQHVDVVKIKNATIIQYQLKGLPETAYQVDIPFAENNIGS 720  
Db 661 ILDGYSISSITIRYKVGKNEHQHVDVVKIKNATIIQYQLKGLPETAYQVDIPFAENNIGS 720  
QY 721 SNPAFSEHVLVTLPESQAPADLGGGKMLLTAIILGSAGMTCLTVLLAFLIILQKLANVQR 780  
Db 721 SNPAFSEHVLVTLPESQAPADLGGGKMLLTAIILGSAGMTCLTVLLAFLIILQKLANVQR 780  
QY 781 MAQAFQNVREEPAVQNSGTLALNRKVNPNPTIYPVLDWMDIKFQDVGIGNFGQVLK 840  
Db 781 MAQAFQNVREEPAVQNSGTLALNRKVNPNPTIYPVLDWMDIKFQDVGIGNFGQVLK 840  
QY 841 ARIKQDGLMDAAIKRMKEYASKDHRDPAGELEVLCCKLGHHPNIINLLGACEHGYLYL 900  
Db 841 ARIKQDGLMDAAIKRMKEYASKDHRDPAGELEVLCCKLGHHPNIINLLGACEHGYLYL 900  
QY 901 ALEYAPHGNLLDPLKRSRVLETPAPAIANSTASTLSSQOLLHFAADVARGMDYLSQKQF 960  
Db 901 ALEYAPHGNLLDPLKRSRVLETPAPAIANSTASTLSSQOLLHFAADVARGMDYLSQKQF 960  
QY 961 IHRDLAARNILGVENVAKIADFGLSRGQEVVYKKTGMGLPVRWMAIESLNSYVYTTNSD 1020  
Db 961 IHRDLAARNILGVENVAKIADFGLSRGQEVVYKKTGMGLPVRWMAIESLNSYVYTTNSD 1020  
QY 1021 VMSYGVLLWEIVSLGTPYCGMTCAELYEKLPGYRLEKPLNCDDEVYDLMQCWREKPY 1080  
Db 1021 VMSYGVLLWEIVSLGTPYCGMTCAELYEKLPGYRLEKPLNCDDEVYDLMQCWREKPY 1080  
QY 1081 ERPSFAQLVSLNRLMBERKTVNTTLYEKFTYAGIDCSAEAA 1124  
Db 1081 ERPSFAQLVSLNRLMBERKTVNTTLYEKFTYAGIDCSAEAA 1124

RESULT 9

US-10-223-081-82  
; Sequence 82, Application US/10223081  
; Publication NO. US2003018686A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Marsters, Scot A.  
; APPLICANT: Pan, James  
; APPLICANT: Stephan, Jean-Philippe F.  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Ye, Weilan  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; TITLE OF INVENTION: TREATMENT OF DISORDERS INVOLVING ANGIOGENESIS  
; FILE REFERENCE: P3235P1C7  
; CURRENT APPLICATION NUMBER: US/10/223,081  
; CURRENT FILING DATE: 2002-08-16  
; PRIOR APPLICATION NUMBER: US 10/081,056  
; PRIOR FILING DATE: 2002-02-20  
; PRIOR APPLICATION NUMBER: US 60/213,637  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: US 60/219,556  
; PRIOR FILING DATE: 2000-07-20  
; PRIOR APPLICATION NUMBER: US 60/220,624  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: US 60/220,664  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: PCT/US00/20710  
; PRIOR FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: US 60/222,695  
; PRIOR FILING DATE: 2000-08-02  
; PRIOR APPLICATION NUMBER: US 09/643,657

; PRIOR FILING DATE: 2000-08-17  
; PRIOR APPLICATION NUMBER: PCT/US00/23522  
; PRIOR FILING DATE: 2000-08-23  
; PRIOR APPLICATION NUMBER: PCT/US00/23328  
; PRIOR FILING DATE: 2000-08-24  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 383  
; SEQ ID NO 82  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-223-081-82  
  
Query Match 100.0%; Score 6013; DB 4; Length 1124;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MDLSASLVLCGVSLLSGTVEGAMDILINSIPLVSDAETSLTCIASGRWPEPIITGRD 60  
Db 1 MDLSASLVLCGVSLLSGTVEGAMDILINSIPLVSDAETSLTCIASGRWPEPIITGRD 60  
QY 61 FEALNNQHQPLEVTDVTRWAKKVVWKRKASKINGAYFCSEGRVGEAIRIRTMKMRQ 120  
Db 61 FEALNNQHQPLEVTDVTRWAKKVVWKRKASKINGAYFCSEGRVGEAIRIRTMKMRQ 120  
QY 121 QASFLPATLTMTVDKGDNNISFKKVLIKBEDAVIYKNGSFTHSVPRHVPDILEVHLPH 180  
Db 121 QASFLPATLTMTVDKGDNNISFKKVLIKBEDAVIYKNGSFTHSVPRHVPDILEVHLPH 180  
QY 181 AQPQDAGVYSARYIGGNLFTSAFTLIVRRCAQKMGPECNHLCCTACMNGVCHEDTGEC 240  
Db 181 AQPQDAGVYSARYIGGNLFTSAFTLIVRRCAQKMGPECNHLCCTACMNGVCHEDTGEC 240  
QY 241 ICPPGFMGRTCEKACELHTFGRTCKERGSGQCKSYVFCCLPDPYGCSCATGKGLQCN 300  
Db 241 ICPPGFMGRTCEKACELHTFGRTCKERGSGQCKSYVFCCLPDPYGCSCATGKGLQCN 300  
QY 301 ACHPGFYGPDCCKLRCSNNGEMCDRFGQCLCSPGWQGLQCEGREGIPRMTPKIVDLPD 360  
Db 301 ACHPGFYGPDCCKLRCSNNGEMCDRFGQCLCSPGWQGLQCEGREGIPRMTPKIVDLPD 360  
QY 361 VNSGKFNPICKASGHPLEPTNEEMTLVKDGTVLHPKDNHTDHFSAIETIHRILPPDSG 420  
Db 361 VNSGKFNPICKASGHPLEPTNEEMTLVKDGTVLHPKDNHTDHFSAIETIHRILPPDSG 420  
QY 421 VMVCSVNTVAGNVEKPFNISVKVLPKLPANPNVDTGNFVAVINISSEPYFGDGPISK 480  
Db 421 VMVCSVNTVAGNVEKPFNISVKVLPKLPANPNVDTGNFVAVINISSEPYFGDGPISK 480  
QY 481 LLYKPNHYEAWQHIQVTNEIVTLNLYLSPRTYELCVQLVRRGEGEGHPGVRRTTAS 540  
Db 481 LLYKPNHYEAWQHIQVTNEIVTLNLYLSPRTYELCVQLVRRGEGEGHPGVRRTTAS 540  
QY 541 IGLPPRGNLNLLPKSQTTLNLTWQIPFSSSEDDFVEVERRSVQKSDQONIKVPGNLT 600  
Db 541 IGLPPRGNLNLLPKSQTTLNLTWQIPFSSSEDDFVEVERRSVQKSDQONIKVPGNLT 600  
QY 601 LNNLHPREQYVVRARVNTKAGWSEDLTAWTSLDILPPPENIKISNITHSSAVISWT 660  
Db 601 LNNLHPREQYVVRARVNTKAGWSEDLTAWTSLDILPPPENIKISNITHSSAVISWT 660  
QY 661 ILDGYSISSITIRYKVGKNEHQHVDVVKIKNATIIQYQLKGLPETAYQVDIPFAENNIGS 720  
Db 661 ILDGYSISSITIRYKVGKNEHQHVDVVKIKNATIIQYQLKGLPETAYQVDIPFAENNIGS 720  
QY 721 SNPAFSEHVLVTLPESQAPADLGGGKMLLTAIILGSAGMTCLTVLLAFLIILQKLANVQR 780  
Db 721 SNPAFSEHVLVTLPESQAPADLGGGKMLLTAIILGSAGMTCLTVLLAFLIILQKLANVQR 780  
QY 781 MAQAFQNVREEPAVQNSGTLALNRKVNPNPTIYPVLDWMDIKFQDVGIGNFGQVLK 840  
Db 781 MAQAFQNVREEPAVQNSGTLALNRKVNPNPTIYPVLDWMDIKFQDVGIGNFGQVLK 840







QY 1021 VMSYGVLLWEIVSLGTPYCGMTCAELYEKLPQGRLEKPLNCDDEVDLMRQCWREKPY 1080  
DB 1021 VMSYGVLLWEIVSLGTPYCGMTCAELYEKLPQGRLEKPLNCDDEVDLMRQCWREKPY 1080  
QY 1081 ERPSAQILVSLNRMLEERKTYVNTTLYEKFTYAGIDCSAEAA 1124  
DB 1081 ERPSAQILVSLNRMLEERKTYVNTTLYEKFTYAGIDCSAEAA 1124  
RESULT 12  
US-10-305-654-82  
; Sequence 82, Application US/10305654  
; Publication No. US20030224964A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Gerber, Hans-Peter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Goddard, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Marstera, Scot A.  
; APPLICANT: Pan, J.  
; APPLICANT: Paoni, N. F.  
; APPLICANT: Stephan, J.-P. F.  
; APPLICANT: Watanabe, C.K.  
; APPLICANT: Wood, W.I.  
; APPLICANT: Williams, P.M.  
; APPLICANT: Ye, Weilan  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; TITLE OF INVENTION: TREATMENT OF DISORDERS INVOLVING ANGIOGENESIS  
; FILE REFERENCE: P3235R1C1  
; CURRENT APPLICATION NUMBER: US/10/305,654  
; CURRENT FILING DATE: 2002-11-26  
; NUMBER OF SEQ ID NOS: 383  
; SEQ ID NO 82  
; LENGTH: 1124  
; TYPE: PRN  
; ORGANISM: Homosapiens  
US-10-305-654-82  
Query Match 100.0%; Score 6013; DB 4; Length 1124;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDLSASLVLCGVSLLSGTVEGAMDILILINSIPLVSDAETSLTCTASGRWPHETITGRD 60  
DB 1 MDLSASLVLCGVSLLSGTVEGAMDILILINSIPLVSDAETSLTCTASGRWPHETITGRD 60  
QY 61 FEALANQHDPLEVTQDVTREWAKVWVKREXASKINGAYFCEGRVGRGEAIRITWKMQRQ 120  
DB 61 FEALANQHDPLEVTQDVTREWAKVWVKREXASKINGAYFCEGRVGRGEAIRITWKMQRQ 120  
QY 121 QASFLPATLTMTVDKGNVNIISFKVLIKEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
DB 121 QASFLPATLTMTVDKGNVNIISFKVLIKEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180  
QY 181 AQPDAGVYSARYIGNLFTSAFTRLIVRCEAQWGPCNHLCTACWNGVCHEDTGE 240  
DB 181 AQPDAGVYSARYIGNLFTSAFTRLIVRCEAQWGPCNHLCTACWNGVCHEDTGE 240  
QY 241 ICPDGMGTCTCKACELHFTFGRTCKERCSCGQCKSYVFCPLDPYGCSCATWGKLCQNE 300  
DB 241 ICPDGMGTCTCKACELHFTFGRTCKERCSCGQCKSYVFCPLDPYGCSCATWGKLCQNE 300  
QY 301 ACHPGFYGPDCKLRCSNNGCNDPFGQGLCSFGWQGLQCEREGIPRMTPKIVLDPDHE 360  
DB 301 ACHPGFYGPDCKLRCSNNGCNDPFGQGLCSFGWQGLQCEREGIPRMTPKIVLDPDHE 360  
QY 361 VNSGFNPKICKASGWPLPNEEMTLVKPDGTVHLKPDFNHTDHFSAFTIIRIILPPDSG 420  
DB 361 VNSGFNPKICKASGWPLPNEEMTLVKPDGTVHLKPDFNHTDHFSAFTIIRIILPPDSG 420

QY 421 VVCSVNTVAGVKEKPFNISVKVLPKPLNAPNVIDTGNFAVINISSBSPYFGDGIKSKK 480  
DB 421 VVCSVNTVAGVKEKPFNISVKVLPKPLNAPNVIDTGNFAVINISSBSPYFGDGIKSKK 480  
QY 481 LLYKPVNHYEAWOHIQVNTNEIVTLNLYLEPRTEYELCVOLVRRGEGEGHPGVRPFTTAS 540  
DB 481 LLYKPVNHYEAWOHIQVNTNEIVTLNLYLEPRTEYELCVOLVRRGEGEGHPGVRPFTTAS 540  
QY 541 IGLPPRGILNLLPKSQTTNLNTWQPIFPSSSEDDFVVEVERRSVQKSDQONIKVPGNLTSV 600  
DB 541 IGLPPRGILNLLPKSQTTNLNTWQPIFPSSSEDDFVVEVERRSVQKSDQONIKVPGNLTSV 600  
QY 601 LLNLHPREQYVVRARVNTKAGWSEDLTAWTILSDILPPQENIKISNITHSSAVISWT 660  
DB 601 LLNLHPREQYVVRARVNTKAGWSEDLTAWTILSDILPPQENIKISNITHSSAVISWT 660  
QY 661 ILDGYSISSITIRYKVGKNEQDQVVKIKNATIIQYOLKGLEPETAYQVDIFAENNIGS 720  
DB 661 ILDGYSISSITIRYKVGKNEQDQVVKIKNATIIQYOLKGLEPETAYQVDIFAENNIGS 720  
QY 721 SNPAFSELVTLPEQAADLGGGKMLLIAIILGSAGMTCLTVLLAFLIILQLKRVQR 780  
DB 721 SNPAFSELVTLPEQAADLGGGKMLLIAIILGSAGMTCLTVLLAFLIILQLKRVQR 780  
QY 781 MAQAFQNVREBPVQFNSGTILALNRKVNNDPPTIYPVLDNDIKFQDVIGSGNFGVLK 840  
DB 781 MAQAFQNVREBPVQFNSGTILALNRKVNNDPPTIYPVLDNDIKFQDVIGSGNFGVLK 840  
QY 841 ARIKDGRLMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNIIILGACEHGRGILYL 900  
DB 841 ARIKDGRLMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNIIILGACEHGRGILYL 900  
QY 901 AIEVAPHGNLLDFLRKSRVLETDPAFAIANSTASTLSSQQLLHFAADVARGMDYLSQKF 960  
DB 901 AIEVAPHGNLLDFLRKSRVLETDPAFAIANSTASTLSSQQLLHFAADVARGMDYLSQKF 960  
QY 961 IHRDLAARNILVGENYVAKIADFLSRGQEVVYKKTGMGRLPVRWMAIESLNTSVYTTNSD 1020  
DB 961 IHRDLAARNILVGENYVAKIADFLSRGQEVVYKKTGMGRLPVRWMAIESLNTSVYTTNSD 1020  
QY 1021 VMSYGVLLWEIVSLGTPYCGMTCAELYEKLPQGRLEKPLNCDDEVDLMRQCWREKPY 1080  
DB 1021 VMSYGVLLWEIVSLGTPYCGMTCAELYEKLPQGRLEKPLNCDDEVDLMRQCWREKPY 1080  
QY 1081 ERPSAQILVSLNRMLEERKTYVNTTLYEKFTYAGIDCSAEAA 1124  
DB 1081 ERPSAQILVSLNRMLEERKTYVNTTLYEKFTYAGIDCSAEAA 1124  
RESULT 12  
US-10-394-322A-64  
; Sequence 64, Application US/10394322A  
; Publication No. US20030232391A1  
; GENERAL INFORMATION:  
; APPLICANT: SUNESIS PHARMACEUTICALS, INC.  
; APPLICANT: Prescott, John C.  
; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS  
; FILE REFERENCE: 39750-0006 US  
; CURRENT APPLICATION NUMBER: US/10/394,322A  
; PRIOR FILING DATE: 2003-03-20  
; PRIOR APPLICATION NUMBER: US 60/366,892  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 64  
; LENGTH: 1124  
; TYPE: PRN  
; ORGANISM: Homo sapiens  
US-10-394-322A-64  
Query Match 100.0%; Score 6013; DB 4; Length 1124;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MDSLASVLCVSLLSGTVGAMDLILINSILPLVSDAETSLTCTASGWRPHEPITIGRD	60
Db	1	MDSLASVLCVSLLSGTVGAMDLILINSILPLVSDAETSLTCTASGWRPHEPITIGRD	60
Qy	61	FEALMNQHDPLEVTQDVTREWAKVVMKREKASKINGAYFCEGVRGEAIRITMKMRQ	120
Db	61	FEALMNQHDPLEVTQDVTREWAKVVMKREKASKINGAYFCEGVRGEAIRITMKMRQ	120
Qy	121	QASFLPATLTMTVDKGNVNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH	180
Db	121	QASFLPATLTMTVDKGNVNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH	180
Qy	181	AQPDAGVYSARYICGNLFTSAFTRLIVRRCEAQKWGPCNHLCTACMNGVCHEDTGEC	240
Db	181	AQPDAGVYSARYICGNLFTSAFTRLIVRRCEAQKWGPCNHLCTACMNGVCHEDTGEC	240
Qy	241	ICPPGFMGRTCEKACELHTFGRTCKERCSGQGGCKSYVFCPLDPDYGCSATGWKGLQONE	300
Db	241	ICPPGFMGRTCEKACELHTFGRTCKERCSGQGGCKSYVFCPLDPDYGCSATGWKGLQONE	300
Qy	301	ACHPGFYGPDCKLRCSNNGEMCDRFQGLCSLSPGWQGLQEREGIPRMTPKIVDLPDHIE	360
Db	301	ACHPGFYGPDCKLRCSNNGEMCDRFQGLCSLSPGWQGLQEREGIPRMTPKIVDLPDHIE	360
Qy	361	VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAIFTTHRIILPPDSG	420
Db	361	VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAIFTTHRIILPPDSG	420
Qy	421	VWVCSVNTVAGWKEKPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPYFGDGPIKSKK	480
Db	421	VWVCSVNTVAGWKEKPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPYFGDGPIKSKK	480
Qy	481	LLYKFPVNHAEAWHQHIVQTNIEIVTLNLYLPRTEYELCVQLVRRGECEGHPGVRPFTTAS	540
Db	481	LLYKFPVNHAEAWHQHIVQTNIEIVTLNLYLPRTEYELCVQLVRRGECEGHPGVRPFTTAS	540
Qy	541	IGLPPRGNLPLPKSQTTLNLTWQPIFPSSDDDFVVEVERSVQSDQONIKVGNLTSV	600
Db	541	IGLPPRGNLPLPKSQTTLNLTWQPIFPSSDDDFVVEVERSVQSDQONIKVGNLTSV	600
Qy	601	LLNLNHPREQYVVRVNTKQGEWSEDLTAWTSLDILPPDENIKISNIHSSAVISWT	660
Db	601	LLNLNHPREQYVVRVNTKQGEWSEDLTAWTSLDILPPDENIKISNIHSSAVISWT	660
Qy	661	ILDGYSISSITIRYKVGQKNEQHDVVKIKNATIIQYQLKGLEPETAQVDFIAENNIGS	720
Db	661	ILDGYSISSITIRYKVGQKNEQHDVVKIKNATIIQYQLKGLEPETAQVDFIAENNIGS	720
Qy	721	SNPAFSELVTLPEAOADIGGGKMLLIIILGSAGMTCLTVLLAFILIIQLKRANVQR	780
Db	721	SNPAFSELVTLPEAOADIGGGKMLLIIILGSAGMTCLTVLLAFILIIQLKRANVQR	780
Qy	781	MAQAFQNVREBPVQFNSGTLALNRKVNNDPTIYPVLDWMDIKFQDVI GEGNFGQVLK	840
Db	781	MAQAFQNVREBPVQFNSGTLALNRKVNNDPTIYPVLDWMDIKFQDVI GEGNFGQVLK	840
Qy	841	ARIKKGDLRMDAAIKRMKEYASKDHRDPAGELEVLCXKLGHHNPTINILGACEHRGYLYL	900
Db	841	ARIKKGDLRMDAAIKRMKEYASKDHRDPAGELEVLCXKLGHHNPTINILGACEHRGYLYL	900
Qy	901	AIEYAPHNLJDLFLKRSVLETDPAFANSTASTLSQQLLHFAADVARGMDYLSQKOF	960
Db	901	AIEYAPHNLJDLFLKRSVLETDPAFANSTASTLSQQLLHFAADVARGMDYLSQKOF	960
Qy	961	IHRDLAARNILVGENYAKIADFGLSRQGEVVYKKTGMRLPVRNAIESLNAVSVYTTNSD	1020
Db	961	IHRDLAARNILVGENYAKIADFGLSRQGEVVYKKTGMRLPVRNAIESLNAVSVYTTNSD	1020
Qy	1021	VMSYGVLLWEIVSLGGTYPGCMTCAELEKULPQGYRLEKPLNCDDEVYDLMRQCWREKPY	1080
Db	1021	VMSYGVLLWEIVSLGGTYPGCMTCAELEKULPQGYRLEKPLNCDDEVYDLMRQCWREKPY	1080

Qy	1081	ERPSAQILVSLNRMLEERKTYVNTTLYEKPTYAGIDCSABEAA	1124
Db	1081	ERPSAQILVSLNRMLEERKTYVNTTLYEKPTYAGIDCSABEAA	1124
RESULT 13			
US-10-211-462-147			
; Sequence 147, Application US/10211462			
; Publication No. US20040033495A1			
; GENERAL INFORMATION:			
; APPLICANT: Murray, Richard			
; APPLICANT: Glynn, Richard			
; APPLICANT: Watson, Susan R.			
; APPLICANT: Aziz, Natasha			
; APPLICANT: Eos Biotechnology, Inc.			
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and			
; TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators			
; FILE REFERENCE: 018501-006200US			
; CURRENT APPLICATION NUMBER: US/10/211,462			
; CURRENT FILING DATE: 2003-02-13			
; PRIOR APPLICATION NUMBER: US 09/784,356			
; PRIOR FILING DATE: 2001-02-14			
; PRIOR APPLICATION NUMBER: US 09/791,390			
; PRIOR FILING DATE: 2001-02-22			
; PRIOR APPLICATION NUMBER: US 60/310,025			
; PRIOR FILING DATE: 2001-08-03			
; PRIOR APPLICATION NUMBER: US 60/334,244			
; PRIOR FILING DATE: 2001-11-29			
; NUMBER OF SEQ ID NOS: 230			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 147			
; LENGTH: 1124			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-211-462-147			

Query Match 100.0%; Score 6013; DB 4; Length 1124;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MDSLASVLCVSLLSGTVGAMDLILINSILPLVSDAETSLTCTASGWRPHEPITIGRD	60
Db	1	MDSLASVLCVSLLSGTVGAMDLILINSILPLVSDAETSLTCTASGWRPHEPITIGRD	60
Qy	61	FEALMNQHDPLEVTQDVTREWAKVVMKREKASKINGAYFCEGVRGEAIRITMKMRQ	120
Db	61	FEALMNQHDPLEVTQDVTREWAKVVMKREKASKINGAYFCEGVRGEAIRITMKMRQ	120
Qy	121	QASFLPATLTMTVDKGNVNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH	180
Db	121	QASFLPATLTMTVDKGNVNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH	180
Qy	181	AQPDAGVYSARYICGNLFTSAFTRLIVRRCEAQKWGPCNHLCTACMNGVCHEDTGEC	240
Db	181	AQPDAGVYSARYICGNLFTSAFTRLIVRRCEAQKWGPCNHLCTACMNGVCHEDTGEC	240
Qy	241	ICPPGFMGRTCEKACELHTFGRTCKERCSGQGGCKSYVFCPLDPDYGCSATGWKGLQONE	300
Db	241	ICPPGFMGRTCEKACELHTFGRTCKERCSGQGGCKSYVFCPLDPDYGCSATGWKGLQONE	300
Qy	301	ACHPGFYGPDCKLRCSNNGEMCDRFQGLCSLSPGWQGLQEREGIPRMTPKIVDLPDHIE	360
Db	301	ACHPGFYGPDCKLRCSNNGEMCDRFQGLCSLSPGWQGLQEREGIPRMTPKIVDLPDHIE	360
Qy	361	VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAIFTTHRIILPPDSG	420
Db	361	VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFSAIFTTHRIILPPDSG	420
Qy	421	VWVCSVNTVAGWKEKPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPYFGDGPIKSKK	480
Db	421	VWVCSVNTVAGWKEKPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPYFGDGPIKSKK	480

QY 481 LLYKPVNHYEAWQHIOVNEIVTLNLYBRTYELCVQLVRRGEGBGHPGVRFTTAS 540  
 Db 481 LLYKPVNHYEAWQHIOVNEIVTLNLYBRTYELCVQLVRRGEGBGHPGVRFTTAS 540  
 QY 541 IGLPPRGNLNLPKSGOTTLNLTWQPIFPSSDDFFVEVRRSVQKSDQONIKVPGNLTSV 600  
 Db 541 IGLPPRGNLNLPKSGOTTLNLTWQPIFPSSDDFFVEVRRSVQKSDQONIKVPGNLTSV 600  
 QY 601 LNNLHPREYVVRARVNTKAQWSEDLTAWTSLDILPPQENIKISNITHSSAVISWT 660  
 Db 601 LNNLHPREYVVRARVNTKAQWSEDLTAWTSLDILPPQENIKISNITHSSAVISWT 660  
 QY 661 ILDGYSISITIRYKQGNEDQHDVVKIKNATIIQYQLKGLBPTAYQVDIFAEWNIGS 720  
 Db 661 ILDGYSISITIRYKQGNEDQHDVVKIKNATIIQYQLKGLBPTAYQVDIFAEWNIGS 720  
 QY 721 SNPAFSHELVTLPESQAPADLGGGKMLLIIAGSAGMTCLTVLLAFLIILQKRVNORR 780  
 Db 721 SNPAFSHELVTLPESQAPADLGGGKMLLIIAGSAGMTCLTVLLAFLIILQKRVNORR 780  
 QY 781 MAQAFQNVREEPVAFNSGTLALNRKVKNNPDPTIYPVLDMDNDIKFQDVIGEGNFGVLK 840  
 Db 781 MAQAFQNVREEPVAFNSGTLALNRKVKNNPDPTIYPVLDMDNDIKFQDVIGEGNFGVLK 840  
 QY 841 ARIKXGDLMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNIIINLIGACEHRGYLYL 900  
 Db 841 ARIKXGDLMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNIIINLIGACEHRGYLYL 900  
 QY 901 ALEYAPHGNLDFLRKSRVLETDPAFATANSTASTLSOQLHFAADVARGMDYLSQKOF 960  
 Db 901 ALEYAPHGNLDFLRKSRVLETDPAFATANSTASTLSOQLHFAADVARGMDYLSQKOF 960  
 QY 961 IHRDLAARNILGVENVAKIADFGLSRGQEVVVKKTGMGLPVRWMAIESLNVSVTTNSD 1020  
 Db 961 IHRDLAARNILGVENVAKIADFGLSRGQEVVVKKTGMGLPVRWMAIESLNVSVTTNSD 1020  
 QY 1021 VMSYGVLLWEIVSLGGTFCYGMTCAELVEKLPQGYRLEKPLNCDDEVYDMRQCWREKPY 1080  
 Db 1021 VMSYGVLLWEIVSLGGTFCYGMTCAELVEKLPQGYRLEKPLNCDDEVYDMRQCWREKPY 1080  
 QY 1081 ERPSFAQILVSLNRLMEERKTVNTLYEKFYAGIDCSAEAA 1124  
 Db 1081 ERPSFAQILVSLNRLMEERKTVNTLYEKFYAGIDCSAEAA 1124

RESULT 14

US-10-081-056-82  
 ; Sequence 82, Application US/10081056  
 ; Publication NO. US20040043927A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Ferrara, Napoleone  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Hillan, Kenneth J.  
 ; APPLICANT: Marsters, Scot A.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Pao, Nicholas F.  
 ; APPLICANT: Stephan, Jean-Philippe F.  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Wood, William I.  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Ye, Weilian  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
 ; TREATMENT OF DISORDERS INVOLVING ANGIOGENESIS  
 ; FILE REFERENCE: P3235P1C1  
 ; CURRENT APPLICATION NUMBER: US/10/081,056  
 ; CURRENT FILING DATE: 2002-02-20  
 ; PRIOR APPLICATION NUMBER: PCT/US01/21735  
 ; PRIOR FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: US 60/219,556  
 ; PRIOR FILING DATE: 2000-07-20  
 ; PRIOR APPLICATION NUMBER: US 60/220,624  
 ; PRIOR FILING DATE: 2000-07-25  
 ; PRIOR APPLICATION NUMBER: US 60/220,664  
 ; PRIOR FILING DATE: 2000-07-25  
 ; PRIOR APPLICATION NUMBER: PCT/US00/20710  
 ; PRIOR FILING DATE: 2000-07-28  
 ; PRIOR APPLICATION NUMBER: US 60/222,695  
 ; PRIOR FILING DATE: 2000-08-02  
 ; PRIOR APPLICATION NUMBER: US 09/643,657  
 ; PRIOR FILING DATE: 2000-08-17  
 ; PRIOR APPLICATION NUMBER: PCT/US00/23522  
 ; PRIOR FILING DATE: 2000-08-23  
 ; PRIOR APPLICATION NUMBER: PCT/US00/23328  
 ; PRIOR FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/230,978  
 ; PRIOR FILING DATE: 2000-09-07  
 ; PRIOR APPLICATION NUMBER: US 60/000,000  
 ; PRIOR FILING DATE: 2000-09-15  
 ; PRIOR APPLICATION NUMBER: US 09/664,610  
 ; PRIOR FILING DATE: 2000-09-18  
 ; PRIOR APPLICATION NUMBER: US 09/665,350  
 ; PRIOR FILING DATE: 2000-09-18  
 ; PRIOR APPLICATION NUMBER: US 60/242,922  
 ; PRIOR FILING DATE: 2000-10-24  
 ; PRIOR APPLICATION NUMBER: US 09/709,238  
 ; PRIOR FILING DATE: 2000-11-08  
 ; PRIOR APPLICATION NUMBER: PCT/US00/30952  
 ; PRIOR FILING DATE: 2000-11-08  
 ; PRIOR APPLICATION NUMBER: PCT/US00/30873  
 ; PRIOR FILING DATE: 2000-11-10  
 ; PRIOR APPLICATION NUMBER: PCT/US00/32678  
 ; PRIOR FILING DATE: 2000-12-01  
 ; PRIOR APPLICATION NUMBER: US 09/747,259  
 ; PRIOR FILING DATE: 2000-12-20  
 ; PRIOR APPLICATION NUMBER: PCT/US00/34956  
 ; PRIOR FILING DATE: 2000-12-20  
 ; PRIOR APPLICATION NUMBER: US 09/767,609  
 ; PRIOR FILING DATE: 2001-01-22  
 ; PRIOR APPLICATION NUMBER: US 09/796,498  
 ; PRIOR FILING DATE: 2001-02-28  
 ; PRIOR APPLICATION NUMBER: PCT/US01/06520  
 ; PRIOR FILING DATE: 2001-02-28  
 ; PRIOR APPLICATION NUMBER: PCT/US01/06666  
 ; PRIOR FILING DATE: 2001-03-01  
 ; PRIOR APPLICATION NUMBER: US 09/802,706  
 ; PRIOR FILING DATE: 2001-03-09  
 ; PRIOR APPLICATION NUMBER: US 09/808,689  
 ; PRIOR FILING DATE: 2001-03-14  
 ; PRIOR APPLICATION NUMBER: US 09/816,744  
 ; PRIOR FILING DATE: 2001-03-22  
 ; PRIOR APPLICATION NUMBER: US 09/828,366  
 ; PRIOR FILING DATE: 2001-04-05  
 ; PRIOR APPLICATION NUMBER: US 09/854,208  
 ; PRIOR FILING DATE: 2001-05-10  
 ; PRIOR APPLICATION NUMBER: US 09/854,280  
 ; PRIOR FILING DATE: 2001-05-10  
 ; PRIOR APPLICATION NUMBER: US 09/866,028  
 ; PRIOR FILING DATE: 2001-05-25  
 ; PRIOR APPLICATION NUMBER: US 09/866,034  
 ; PRIOR FILING DATE: 2001-05-25  
 ; PRIOR APPLICATION NUMBER: PCT/US01/17092  
 ; PRIOR FILING DATE: 2001-05-25  
 ; PRIOR APPLICATION NUMBER: US 09/870,574  
 ; PRIOR FILING DATE: 2001-05-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/17443  
 ; PRIOR FILING DATE: 2001-05-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/17800  
 ; PRIOR FILING DATE: 2001-06-01  
 ; PRIOR APPLICATION NUMBER: PCT/US01/19692  
 ; PRIOR FILING DATE: 2001-06-20  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00000

; PRIOR FILING DATE: 2001-06-28  
; NUMBER OF SEQ ID NOS: 383  
; SEQ ID NO 82  
; LENGTH: 1124  
; TYPE: PRT  
; ORGANISM: Homosapiens  
US-10-081-056-82

```
Query Match      100.0%; Score 6013; DB 4; Length 1124;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDSLASLVLCGVSLLSGTVEGAMDILILINSILPLVSDAETSLTCIASGRWHPHEPITIGRD 60
Db 1 MDSLASLVLCGVSLLSGTVEGAMDILILINSILPLVSDAETSLTCIASGRWHPHEPITIGRD 60

Qy 61 FEALMNQHDPLEVTQDVTREWAKKVVWVKREKASKINGAYFCEGRVGEAIRIRTMKMRQ 120
Db 61 FEALMNQHDPLEVTQDVTREWAKKVVWVKREKASKINGAYFCEGRVGEAIRIRTMKMRQ 120

Qy 121 QASFLPATLTMTVDKGNVNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180
Db 121 QASFLPATLTMTVDKGNVNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180

Qy 181 AQPDAGVYSARYIGGNLFTSAFTRLIVRCEAOKWGPCNHLCTACMNGVCHEDTGEC 240
Db 181 AQPDAGVYSARYIGGNLFTSAFTRLIVRCEAOKWGPCNHLCTACMNGVCHEDTGEC 240

Qy 241 ICPPGFMGRGTCCEKACELHTFGRTCKERCSCGGQCKSVYFCLPDPYGCSCATGWKGLQNE 300
Db 241 ICPPGFMGRGTCCEKACELHTFGRTCKERCSCGGQCKSVYFCLPDPYGCSCATGWKGLQNE 300

Qy 301 ACHPGFYGPDCKLRCSNNGEMCDRFQCLCSPGWGLQCEGREGIPRMTPKIVDLPDHE 360
Db 301 ACHPGFYGPDCKLRCSNNGEMCDRFQCLCSPGWGLQCEGREGIPRMTPKIVDLPDHE 360

Qy 361 VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFVAFTHRIILPPDSG 420
Db 361 VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFVAFTHRIILPPDSG 420

Qy 421 VVWCSVNTVAGVKEPFIISVKVLKPKLNAPNVIDTGNFAVINISSEPYFGDGPIKSKK 480
Db 421 VVWCSVNTVAGVKEPFIISVKVLKPKLNAPNVIDTGNFAVINISSEPYFGDGPIKSKK 480

Qy 481 LLYKFPVNYEAWQHIVQVNEIVTLNLPRTYELCVOLVRGECEGHPGVRFFTTAS 540
Db 481 LLYKFPVNYEAWQHIVQVNEIVTLNLPRTYELCVOLVRGECEGHPGVRFFTTAS 540

Qy 541 IGLPPRGGLNLLPKSQTLNLTWQPIFPSSDDDFVEVERRSVQSDQONIKVPGNLTSV 600
Db 541 IGLPPRGGLNLLPKSQTLNLTWQPIFPSSDDDFVEVERRSVQSDQONIKVPGNLTSV 600

Qy 601 LNNLHPREQYVVRVNTKAGWSEDLTAWTSLDILPPQENIKIKNITHSSAVISWT 660
Db 601 LNNLHPREQYVVRVNTKAGWSEDLTAWTSLDILPPQENIKIKNITHSSAVISWT 660

Qy 661 ILDGYSISITIRYKVGKNEHDVVKIKNATIIQYOLKLEPETAQYVDIFAENNTGS 720
Db 661 ILDGYSISITIRYKVGKNEHDVVKIKNATIIQYOLKLEPETAQYVDIFAENNTGS 720

Qy 721 SNPAFSELVTLPEAQAPADLGGGKMLLIALIGSAGMTCLTVLLAFLLIOLKRVANVORR 780
Db 721 SNPAFSELVTLPEAQAPADLGGGKMLLIALIGSAGMTCLTVLLAFLLIOLKRVANVORR 780

Qy 781 MAQAFQNVREBPVAFNSGTLALNRKVNPNDDPTIYPVLDWMDIKFDVIGEGNFGQVLK 840
Db 781 MAQAFQNVREBPVAFNSGTLALNRKVNPNDDPTIYPVLDWMDIKFDVIGEGNFGQVLK 840

Qy 841 ARIKDDGLRMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNLIINLLGACEHRGYLYL 900
Db 841 ARIKDDGLRMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNLIINLLGACEHRGYLYL 900

Qy 901 AIEYAPHGNLLDFLRKSRVLETPAFATANSTASTLSSQQLLHFAADVARGMDYLSQKQF 960
```

```
Db 901 AIEYAPHGNLLDFLRKSRVLETPAFATANSTASTLSSQQLLHFAADVARGMDYLSQKQF 960
Qy 961 IHRDLAARNILVGENYVAKIADFGLSRGQEVVVKKTMGRLPVRWMAIESLNSVYTTNSD 1020
Db 961 IHRDLAARNILVGENYVAKIADFGLSRGQEVVVKKTMGRLPVRWMAIESLNSVYTTNSD 1020

Qy 1021 VWSGVLLWEIVSLGGTPYCGMTCAELYEKLPQGVRLKPLNCCDDDEVVDLMRQCWREKPY 1080
Db 1021 VWSGVLLWEIVSLGGTPYCGMTCAELYEKLPQGVRLKPLNCCDDDEVVDLMRQCWREKPY 1080

Qy 1081 ERPSPAQILVSLNRMLEERKTYVNTTLYEKFTYAGIDCSABEAA 1124
Db 1081 ERPSPAQILVSLNRMLEERKTYVNTTLYEKFTYAGIDCSABEAA 1124

RESULT 15
US-11-019-855-47
; Sequence 47, Application US/11019855
; Publication No. US20050158771A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann-La Roche Inc.
; TITLE OF INVENTION: novel targets for obesity from skeletal muscle
; FILE REFERENCE: case 22314
; CURRENT APPLICATION NUMBER: US/11/019,855
; CURRENT FILING DATE: 2004-12-22
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 47
; LENGTH: 1124
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: TEK tyrosine kinase
; LOCATION: (1)..(1124)
; OTHER INFORMATION: LocuID: 7010
; OTHER INFORMATION: NM_000459
US-11-019-855-47

Query Match      100.0%; Score 6013; DB 6; Length 1124;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDSLASLVLCGVSLLSGTVEGAMDILILINSILPLVSDAETSLTCIASGRWHPHEPITIGRD 60
Db 1 MDSLASLVLCGVSLLSGTVEGAMDILILINSILPLVSDAETSLTCIASGRWHPHEPITIGRD 60

Qy 61 FEALMNQHDPLEVTQDVTREWAKKVVWVKREKASKINGAYFCEGRVGEAIRIRTMKMRQ 120
Db 61 FEALMNQHDPLEVTQDVTREWAKKVVWVKREKASKINGAYFCEGRVGEAIRIRTMKMRQ 120

Qy 121 QASFLPATLTMTVDKGNVNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180
Db 121 QASFLPATLTMTVDKGNVNISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180

Qy 181 AQPDAGVYSARYIGGNLFTSAFTRLIVRCEAOKWGPCNHLCTACMNGVCHEDTGEC 240
Db 181 AQPDAGVYSARYIGGNLFTSAFTRLIVRCEAOKWGPCNHLCTACMNGVCHEDTGEC 240

Qy 241 ICPPGFMGRGTCCEKACELHTFGRTCKERCSCGGQCKSVYFCLPDPYGCSCATGWKGLQNE 300
Db 241 ICPPGFMGRGTCCEKACELHTFGRTCKERCSCGGQCKSVYFCLPDPYGCSCATGWKGLQNE 300

Qy 301 ACHPGFYGPDCKLRCSNNGEMCDRFQCLCSPGWGLQCEGREGIPRMTPKIVDLPDHE 360
Db 301 ACHPGFYGPDCKLRCSNNGEMCDRFQCLCSPGWGLQCEGREGIPRMTPKIVDLPDHE 360

Qy 361 VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFVAFTHRIILPPDSG 420
Db 361 VNSGKFNPICKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFVAFTHRIILPPDSG 420

Qy 421 VVWCSVNTVAGVKEPFIISVKVLKPKLNAPNVIDTGNFAVINISSEPYFGDGPIKSKK 480
Db 421 VVWCSVNTVAGVKEPFIISVKVLKPKLNAPNVIDTGNFAVINISSEPYFGDGPIKSKK 480
```

Db 421 VVWCVNTVAGVKEPFIISVKVLKPLNAPNVIDTGHNFVAVINISSEPFYDGPISKKK 480  
Qy 481 LLYKPVNHYEAMQHIOVTNEIIVTLNLYEPTEYELCVQVLRREGEHGPVRRFTTAS 540  
Db 481 LLYKPVNHYEAMQHIOVTNEIIVTLNLYEPTEYELCVQVLRREGEHGPVRRFTTAS 540  
Qy 541 IGLPPRGINLLPKSQTTLLNLTWQIFPSSSEDDFVEVERRSVQKSDOONIKVPGNLTSV 600  
Db 541 IGLPPRGINLLPKSQTTLLNLTWQIFPSSSEDDFVEVERRSVQKSDOONIKVPGNLTSV 600  
Qy 601 LNNLHPREQYVVRARVNTKAGWSEDTATLSDILPPQENIKISNITHSSAVISWT 660  
Db 601 LNNLHPREQYVVRARVNTKAGWSEDTATLSDILPPQENIKISNITHSSAVISWT 660  
Qy 661 ILDGYSISSITIRYKVOGKNEHQHVDVKIKNATIIQYQKGLPETAYQVDIFAENNIGS 720  
Db 661 ILDGYSISSITIRYKVOGKNEHQHVDVKIKNATIIQYQKGLPETAYQVDIFAENNIGS 720  
Qy 721 SNPAFSHELVTLPESQAPADLGGGKMLLIAILGSAGMTCLTVLLAFLIILQLKRANVORR 780  
Db 721 SNPAFSHELVTLPESQAPADLGGGKMLLIAILGSAGMTCLTVLLAFLIILQLKRANVORR 780  
Qy 781 MAQAFQNVREEPVAFNSGTLALNRKVNNDPTIYPVLDWNDIKFQDVIGEGNFGVLK 840  
Db 781 MAQAFQNVREEPVAFNSGTLALNRKVNNDPTIYPVLDWNDIKFQDVIGEGNFGVLK 840  
Qy 841 ARIKKDGLRMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNIINLIGACEHRGYLYL 900  
Db 841 ARIKKDGLRMDAAIKRMKEYASKDHRDPAGELEVLCKLGHHPNIINLIGACEHRGYLYL 900  
Qy 901 AIEYAPHGNLLDPLRKSRLVETDPAPAIANSTASTLSSQQLLHPAADVARGMDYLSQKQF 960  
Db 901 AIEYAPHGNLLDPLRKSRLVETDPAPAIANSTASTLSSQQLLHPAADVARGMDYLSQKQF 960  
Qy 961 IHRDLAARNILVGENYVAKIADFGLSRGQEVYVKTGMGELPVRWMAIESLNYSVYTTNSD 1020  
Db 961 IHRDLAARNILVGENYVAKIADFGLSRGQEVYVKTGMGELPVRWMAIESLNYSVYTTNSD 1020  
Qy 1021 VMSYGVLLWEIVSLGTPYCGMTCAELYEKLPQGYRLEKPLNCDDEVDLMDRCWREKPY 1080  
Db 1021 VMSYGVLLWEIVSLGTPYCGMTCAELYEKLPQGYRLEKPLNCDDEVDLMDRCWREKPY 1080  
Qy 1081 ERPSPAQILVSLNRMLEBKTYVNTTLYEKFTYAGIDCSAEAAA 1124  
Db 1081 ERPSPAQILVSLNRMLEBKTYVNTTLYEKFTYAGIDCSAEAAA 1124

Search completed: March 14, 2006, 02:03:15  
Job time : 181 secs

**This Page Blank (uspto)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2006, 02:00:23 ; Search time 24 Seconds  
(without alignments)  
1303.600 Million cell updates/sec

Title: US-10-633-742-8  
Perfect score: 6013  
Sequence: 1 MDSLASLVLCVSGVLLSGTV.....FTLYEKFTYAGIDCSABEAA 1124

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 161667 segs, 27834885 residues

Total number of hits satisfying chosen parameters: 161667

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA\_New:\*  
1: /cgn2\_6/prodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/prodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
3: /cgn2\_6/prodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
4: /cgn2\_6/prodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
5: /cgn2\_6/prodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
6: /cgn2\_6/prodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
7: /cgn2\_6/prodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
8: /cgn2\_6/prodata/1/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6013	100.0	1124	7	US-11-049-536-162 Sequence 162, App
2	2527	42.0	1138	7	US-11-049-536-2 Sequence 2, Appli
3	696.5	11.6	822	7	US-11-183-567A-2 Sequence 2, Appli
4	659	11.0	822	7	US-11-072-175-254 Sequence 254, App
5	646.5	10.8	258	6	US-10-877-346-75 Sequence 75, Appl
6	633	10.5	310	7	US-11-092-168-9 Sequence 9, Appli
7	612	10.2	985	7	US-11-113-424-61 Sequence 61, Appl
8	611.5	10.2	983	7	US-11-113-424-59 Sequence 59, Appl
9	611.5	10.2	983	7	US-11-203-251A-78 Sequence 78, Appl
10	610	10.1	972	7	US-11-177-894-17 Sequence 17, Appl
11	608	10.1	976	7	US-11-148-770-31 Sequence 31, Appl
12	608	10.1	976	7	US-11-177-894-15 Sequence 15, Appl
13	608	10.1	976	7	US-11-177-894-16 Sequence 16, Appl
14	608	10.1	976	7	US-11-177-894-18 Sequence 18, Appl
15	608	10.1	976	7	US-11-177-894-20 Sequence 20, Appl
16	608	10.1	976	7	US-11-177-894-21 Sequence 21, Appl
17	608	10.1	1338	7	US-11-109-156-23 Sequence 23, Appl
18	608	10.1	1338	7	US-11-075-047A-2 Sequence 2, Appli
19	606	10.1	976	7	US-11-177-894-19 Sequence 19, Appl
20	606	10.1	1338	6	US-10-821-234-1622 Sequence 1622, Ap
21	601	10.0	986	7	US-11-203-251A-80 Sequence 80, Appl
22	593.5	9.9	972	6	US-10-821-234-1587 Sequence 1587, Ap
23	592.5	9.9	976	6	US-10-966-483-2 Sequence 2, Appli
24	592.5	9.9	976	6	US-10-511-273-1 Sequence 1, Appli
25	592.5	9.9	976	7	US-11-233-796-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-11-049-536-162  
; Sequence 162, Application US/11049536  
; Publication No. US20060024297A1  
; GENERAL INFORMATION:  
; APPLICANT: Wood, Clive R.  
; APPLICANT: Dransfield, Daniel T.  
; APPLICANT: Pieters, Henk  
; APPLICANT: Hoet, Rene  
; APPLICANT: Hufton, Simon E.  
; TITLE OF INVENTION: TIE COMPLEX BINDING PROTEINS  
; FILE REFERENCE: 10280-128001  
; CURRENT APPLICATION NUMBER: US/11/049,536  
; CURRENT FILING DATE: 2005-02-02  
; PRIOR APPLICATION NUMBER: US 10/916,840  
; PRIOR FILING DATE: 2004-08-12  
; PRIOR APPLICATION NUMBER: US 60/494,713  
; PRIOR FILING DATE: 2003-08-12  
; NUMBER OF SEQ ID NOS: 721  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 162  
; LENGTH: 1124  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-049-536-162

Query Match	100.0%	Score 6013;	DB 7;	Length 1124;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1124;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MDSLASLVLCVSGVLLSGTVGAGMDLILINSPLVSDAETSLTCTASGRWPHETIGRD	60	
Db	1	MDSLASLVLCVSGVLLSGTVGAGMDLILINSPLVSDAETSLTCTASGRWPHETIGRD	60	
QY	61	FEALMNQHDPLEVTDVYTRWAKVWVKREKASKINGAYFCEGVRGEAIRITMKMQ	120	
Db	61	FEALMNQHDPLEVTDVYTRWAKVWVKREKASKINGAYFCEGVRGEAIRITMKMQ	120	
QY	121	QASFLPATLTMTVDKGDVNIISFKKVLKEEDAVYKNGSFTHSVPRHEVPDILEVHLPH	180	
Db	121	QASFLPATLTMTVDKGDVNIISFKKVLKEEDAVYKNGSFTHSVPRHEVPDILEVHLPH	180	
QY	181	AQPDAGVYSARYIGGNLFTSAFTRLIVRCEAQKMGPECNHLCTACMNGVCHEDTGEC	240	
Db	181	AQPDAGVYSARYIGGNLFTSAFTRLIVRCEAQKMGPECNHLCTACMNGVCHEDTGEC	240	
QY	241	ICPPQFMGRTCEKACELHTFTGRTCHERGSGQGCSSYVFCFLPDPYGCSCATGWKGLQNE	300	

```
Db 241 ICPPGFMGRTEKACELHTFGRTCKERCSCGQBGCKSYVFCPLDPDPYGCSCATGWKGLQCNE 300
Qy 301 ACHPGFYGPDCKLRCSNNGEMCDRFOCLCSPGHGLQCEREGIPRTPKIVLDPDHE 360
Db 301 ACHPGFYGPDCKLRCSNNGEMCDRFOCLCSPGHGLQCEREGIPRTPKIVLDPDHE 360
Qy 361 VNSGKFNPIKASGWPLPNEEMTLVKPDGTVLHPKDFNHTDHFSAIETHRIILPPDSG 420
Db 361 VNSGKFNPIKASGWPLPNEEMTLVKPDGTVLHPKDFNHTDHFSAIETHRIILPPDSG 420
Qy 421 VVWCSVNTVAGWVEKPFNISVKVLKPLNAPNVIDTGNFAVINISSEPPYFGDGPISKK 480
Db 421 VVWCSVNTVAGWVEKPFNISVKVLKPLNAPNVIDTGNFAVINISSEPPYFGDGPISKK 480
Qy 481 LLYKPNHYEAWQHIOVTNEIYTLNLYLPRTEYELCVQLVRGEGEHGPVRRFTTAS 540
Db 481 LLYKPNHYEAWQHIOVTNEIYTLNLYLPRTEYELCVQLVRGEGEHGPVRRFTTAS 540
Qy 541 IGLPPRGLNLLPKSOTTLNLTWQPIFSSSEDDFVVEVERRSVQKSDOONIKVPGNLTSV 600
Db 541 IGLPPRGLNLLPKSOTTLNLTWQPIFSSSEDDFVVEVERRSVQKSDOONIKVPGNLTSV 600
Qy 601 LNNLHPREYQYVVRARVNTKAGWSEDLTAWTSDILPPQENIKISNITHSSAVISWT 660
Db 601 LNNLHPREYQYVVRARVNTKAGWSEDLTAWTSDILPPQENIKISNITHSSAVISWT 660
Qy 661 ILDGYSISSITIRYKVGKNEHDVVDVKIKNATIIYOQLKGLEPETAQVDIFAENNIGS 720
Db 661 ILDGYSISSITIRYKVGKNEHDVVDVKIKNATIIYOQLKGLEPETAQVDIFAENNIGS 720
Qy 721 SNPAFSEHVLTPESQAPADLGGGKMLIAILGSAGMTCCLTVLLAFLIILQKLANVORR 780
Db 721 SNPAFSEHVLTPESQAPADLGGGKMLIAILGSAGMTCCLTVLLAFLIILQKLANVORR 780
Qy 781 MAQAFQNVREEPAVQFNSGTLALNRKVNXPDPPTIYPVLDWMDIKFQDVIGEGNFQVLK 840
Db 781 MAQAFQNVREEPAVQFNSGTLALNRKVNXPDPPTIYPVLDWMDIKFQDVIGEGNFQVLK 840
Qy 841 ARIKDDGLRMDAAIKRMKEYASKDHRDPAGLEVLCKLGHHPNIIINLLGACEHGGYLYL 900
Db 841 ARIKDDGLRMDAAIKRMKEYASKDHRDPAGLEVLCKLGHHPNIIINLLGACEHGGYLYL 900
Qy 901 ALEYAPHGNLLDFLRKSRVLETDPAFANSTASTLSSQQLHFAADVARGMDYLSQKF 960
Db 901 ALEYAPHGNLLDFLRKSRVLETDPAFANSTASTLSSQQLHFAADVARGMDYLSQKF 960
Qy 961 IHRDLAARNILGENYVAKIADFGLSRGQEVYVKTMTGRLPVRWMAIESLNYSVYTTNSD 1020
Db 961 IHRDLAARNILGENYVAKIADFGLSRGQEVYVKTMTGRLPVRWMAIESLNYSVYTTNSD 1020
Qy 1021 VWSYGVLLWEIVSLGTPYCGMTCALYEKLPQGYRLEKPLNCDDDEVYDLMRQCWREKPY 1080
Db 1021 VWSYGVLLWEIVSLGTPYCGMTCALYEKLPQGYRLEKPLNCDDDEVYDLMRQCWREKPY 1080
Qy 1081 ERPSFAQILVSLNRLMEERKTVNTLTKFYAGIDCSAEAAA 1124
Db 1081 ERPSFAQILVSLNRLMEERKTVNTLTKFYAGIDCSAEAAA 1124
```

## RESULT 2

```
US-11-049-536-2
; Sequence 2. Application US/11049536
; Publication No. US20060024297A1
; GENERAL INFORMATION:
; APPLICANT: Wood, Clive R.
; APPLICANT: Dransfield, Daniel T.
; APPLICANT: Pieters, Henk
; APPLICANT: Hoet, Rene
; APPLICANT: Hufton, Simon E.
; TITLE OF INVENTION: TIE COMPLEX BINDING PROTEINS
; FILE REFERENCE: 10280-128001
; CURRENT APPLICATION NUMBER: US/11/049,536
; CURRENT FILING DATE: 2005-02-02
```

```
; PRIOR APPLICATION NUMBER: US 10/916,840
; PRIOR FILING DATE: 2004-08-12
; PRIOR APPLICATION NUMBER: US 60/494,713
; PRIOR FILING DATE: 2003-08-12
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1138
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-049-536-2

Query Match 42.0%; Score 2527; DB 7; Length 1138;
Best Local Similarity 46.7%; Pred. No. 1,1e-162;
Matches 541; Conservative 165; Mismatches 372; Indels 80; Gaps 27;

Qy 14 LLLSGTVEGAMDLLILNSLPLVSDAETSITCI-----ASGWRPHEPITIGRDFAL 64
Db 13 LFLASHVGAADVLTLLANLRLTDPQRFELTCVSGEAGAGRGSDAWGP--PLLLEKDDRIV 70
Qy 65 MNQHQDPLEVTQDVTRWAKKVVWKR-EKASKINGAYCEGVRGEAIRIRTMKMRQOAS 123
Db 71 RTPPGPPLR-----LARNGSHQVTLRGFSKPSDLVGVFSCVCGAGARRTRVIYVHNSPGAH 126
Qy 124 FLPATLTMTVDKGNVNISFKKVLIKERDAVIYKNGSFIHSVPRHEVPD---ILEVHLPH 180
Db 127 LLPDKVHTVNGDVTAVLSARVHEKEQTDVITWKSNGSYFYTLDDWHEAQDGRFLQ--LPN 184
Qy 181 AQPDAGVYSARYIGGNLFTSAFTLIVRCEAQKWGPCNHLCTACMNNGVCHEDTQEC 240
Db 185 VQPPSSGIYSATYLEASPLGSAFFRLIVRGCGAGRWGPGCTKECPCLHGGVCHDHGEC 244
Qy 241 ICPPGFMGRTEKACELHTFGRTCKERCSCGQBGCKSYVFCPLDPDPYGCSCATGWKGLQCNE 300
Db 245 VCPPEFTGTRCEQACREGRFQCGQEQCPGISGCRGLTFCPLDPDPYGCSCGSGWRGSCQE 304
Qy 301 ACHPGFYGPDCKLRCSNNGEMCDRFOCLCSPGHGLQCEREGIPRTPKIVLDPDHE 360
Db 305 ACAPGHFADGCLQCQCGGTCDRFSGCVSPGSMHGVHCEKSD---RIPQLNWAASELE 361
Qy 361 VNSGKFNPI-KASGWPLPNEEMTLVKPDGTVLHPKDFNHTDHFSAIETHRIILPPDS 419
Db 362 FNLETMPRINCAAGNPPVVRGSIELRKPDPGTLLSTKAIIVEPEKTTAEFEVPRVLVADS 421
Qy 420 GVVWCSVNTVAGWVEKPFNISVKVLKPLNAPNVIDTGNFAVIN--ISSEPPYFGDGP 477
Db 422 GFWECRVSTSGQDSRRFRKVNVPVFLAAPRLITKOSRQLVWSPLVS---FSGDGPIS 478
Qy 478 SKKLLYKPNHYEAWQHIOV-TNEIVTLNLYLPRTEYELCVQLVRGEGEHGPVRRF 536
Db 479 TVRLHYRPQDSTMDWSTIVDPSENVTILNLRPKTGYSVRQLSRPGEHGGEGANGPTLM 538
Qy 537 TTASIGLPPPLGLNLLP-----KSQTLNLTWQ-PFPPSS--EDDPYVEVERRSVOKS 586
Db 539 TT--DCPEPL--LQPWLEGHVEGTDRLRVWSLSPLVPGPLVGDGFLRLMDGTGRQE 592
Qy 587 DQONIKVPGNLTSLVLLNHLNHPREYQYVVRARV-NTKAQCEWSEDLTAWTSDILPPQENI 645
Db 593 RRENVSPPQARTA-LLTGLTGTGTHYQLDVQYHCTLLGPASPPAHV-LLPSPGPPAPRHL 650
Qy 646 KISNITHSSAVISW---TILDGYSISSITIRYKVGKNEQDQ---HVDVKIKNATIIYOQL 699
Db 651 HAQALSDSEIQLTWKHPEALPG-PISKYVVEVQVAGGAGDPLWIDVDPRPEETSTII---- 705
Qy 700 KGLEPETAQVDIFA-----ENNIGSNPAFSEHVLTPESQAPADLGGKML 746
Db 706 RGLNASTRYLFRMRASIQGLGDWSNVTVEESTLGNLQAEQ----PVQESRA-AEEGLDQ 760
Qy 747 LLJATILGSAGMTCCLTVLLAFLIILQKLANVORRMAQAFON-VREEPAVQFNSGTLALNR 805
Db 761 LILAVVGSVATCLTLLAALLTLVCIRKSCLRHRRRTFTYQSGSGEETILQSSGGLTLTR 820
Qy 806 KVKNPNPDTIYPVLDWMDIKFQDVIGEGNFQVLKARIKDDGLRMDAAIKRMKEYASKD 865
```



Db 821 RPKLQPEPLSYVLWEDITFDLIGEGNFGVIRAMIKKQGLKNAALIKMLKEYASEND 880  
QY 866 HRDFAGELEVLCKLGHHPNIINLLGACERHGYLYLAIEYAPHGNLLDFLRKSRVLETPA 925  
Db 881 HRDFAGELEVLCKLGHHPNIINLLGACKNRGYLYLAIEYAPVGNLLDFLRKSRVLETPA 940  
QY 926 FAIANSTASTLSSQQLLHPAADVARGMOLSKQFTHRDLAARNLILVGENYVAKIADPGL 985  
Db 941 FAREHGTASTLSSRQLLRFPASDAANGMQLSEKQFTHRDLAARNLILVGENLASKIADPGL 1000  
QY 986 SRGQVYVYKTKGRPLPVRMALESNLNSVYTTNSDVMSYGVLLWEIVSLGGTPYCGMTCA 1045  
Db 1001 SRGEEVYVYKTKGRPLPVRMALESNLNSVYTTKSDVMSYGVLLWEIVSLGGTPYCGMTCA 1060  
QY 1046 ELYEKLPOGYRLEKPLNCDDEYVDMQCRWKPYRPSFAQILVSLNRLMEERKTYVNT 1105  
Db 1061 ELYEKLPOGYRMEQPRNCDDEYVDMQCRWRPPIERPPFAQILQGLGMLERKAYVNM 1120  
QY 1106 TLYEKFYVAGIDCSABEA 1123  
Db 1121 SLFENFTVAGIDATAEEA 1138  
RESULT 3  
US-11-183-567A-2  
; Sequence 2, Application US/11183567A  
; Publication No. US20060019296A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhou, Ming-Ming  
; APPLICANT: Goldfarb, Mitchell  
; TITLE OF INVENTION: Methods of Identifying Modulators of the  
; FILE REFERENCE: 2459-1-002NCON  
; CURRENT APPLICATION NUMBER: US/11/183,567A  
; CURRENT FILING DATE: 2005-07-18  
; PRIOR APPLICATION NUMBER: 09/757,415  
; PRIOR FILING DATE: 2001-01-09  
; PRIOR APPLICATION NUMBER: 60/175,867  
; PRIOR FILING DATE: 2000-01-12  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 822  
; TYPE: PRP  
; ORGANISM: Mus musculus  
US-11-183-567A-2  
Query Match 11.6%; Score 696.5; DB 7; Length 822;  
Best Local Similarity 27.4%; Pred. No. 3.1e-39;  
Matches 226; Conservative 132; Mismatches 297; Indels 169; Gaps 24;  
QY 371 KASGHPLPTNEWTAKP-----DGTVLHPKQPNHHTDHFSAIF 409  
Db 30 QAQMGVPVEVESLIVHPGDLQLRCRDDVQSNLWRDGVQL--VBSNRT-RITGEEV 86  
QY 410 TTHRIPLPDSGVVSVNTVAGMKEKPFNISVK-VLPKPLNAPNVIDTGHNFANINISSE 468  
Db 87 EYRDSIPADSGLYACVTSFSSDSTTFYFVNVSDALPSSDDDDDDSS-----SSE 137  
QY 469 PVFGDPIKSKLLKYPNVNHWBAHQIQTNEIVTLNLEPTREYELCVQLVRRGEGEG 528  
Db 138 EKETDNTKPNR----RPVAPY--W-----TSPEKMEKKLHAPPAKTVKFKCPSSG 182  
QY 529 HPGVRRFTTASIGLPPPGNLNLPKSOPTLNTWQPIPPSSDDF--YVEVERSVQKS 586  
Db 183 TNPPTLRLWLNKGEKPKDHRIGYKRYATWSIIMDSVVPDKNYTCIVENEYSINHT 242  
QY 587 DQONT-----KVPGNLTSVLLNNL-----HPREQYVVRARVNTKAQOE 624  
Db 243 YQLDVVERSPHPILOAGLPANETVALGSNVEFMCKVYSDPQPHIQWLKHIEVNGSKIG- 301  
QY 625 WSEDITAWTSLDILPPQENIKISNTHSSAVISWTLTDGYSISSITIRYKVOGRNEQOH 684

Db 302 -----PDNLPHYQILKTAGV-----NTTDXE 322  
QY 685 VDVKIKNATIIYQYQLKLEPETAYQVDIFAENNIG-SNPAPFSHELVTLPESQAPADLGG 743  
Db 323 MEV-----LHLRNVSPEDAGBYTCLAGNSIGLSHSAWLTVLEALERPA---VMT 370  
QY 744 GKMLILAILGAGMTCLTVLLAFILLOI-----KQANVORRMA-----QAFQNV 789  
Db 371 SPYLEIIIIYCTGAPLISCMGLSVIIYKMKSGTKKSDFHQSMAVHKLAKSIPLRRQVTVS 430  
QY 790 BEPAVQFNSGTALNRKVKNNPDPITYPV-----LDWNDIKFQDVLGEGNF 835  
Db 431 ADSSASMSNGVL-LVRPSRLSSGTPMAGVSEYELPDPWELPRDLVLVLPKPLGEGCF 489  
QY 836 GOVLKARIKKDGLRMD-----AAIKRMKEYASKDDHRDFAGELEVLCKLGHHPNIINL 888  
Db 490 GOVULA--EALGDKDKPNRVTKVAVKMLKSDATEKDLSDLISEMMKMKWICKHNIINL 547  
QY 889 LGACERHGYLYLAIEYAPHGNLLDFLRKSRVLETPAFAIANSTASTLSSQQLLHPAADV 948  
Db 548 LGACTQDGPVLYVIEYASKNRLREYLQARRPPGLECYCYNPSHPNBEQLSSKDLVSCAYQV 607  
QY 949 ARGMDYLSQKQFIHRDLAARNILVGENYVAKIADFGLSRG---QEVYVYKTKGRPLPVRM 1005  
Db 608 ARGMEYLASKKCIHRDLAARNVLTEDNMVKIADFGLARDIHHIDYIKTKTNGRLPVRM 667  
QY 1006 AIESLNYSVYTTNSDVMSYGVLLWEIVSLGGTPYCGMTCAEYKLPQGYRLEKPLNCD 1065  
Db 668 APEALFDRIIYHQSDVMSYGVLLWEIVSLGGTPYCGVPEELFKLLKEGHRMDKPSNCTN 727  
QY 1066 EYVDMQCRWKPYRPSFAQILVSLNRM--LEERKTYVNTTL 1107  
Db 728 ELYMMRDCWHAVPQSQRPTFKQLVREDLDRIVALTSSQEYLDLSI 771  
RESULT 4  
US-11-072-175-254  
; Sequence 254, Application US/11072175  
; Publication No. US20060029944A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT  
; INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR  
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR  
; FILE REFERENCE: D0273A CIP  
; CURRENT APPLICATION NUMBER: US/11/072,175  
; CURRENT FILING DATE: 2005-03-05  
; PRIOR APPLICATION NUMBER: US 60/406,385  
; PRIOR FILING DATE: 2002-08-27  
; PRIOR APPLICATION NUMBER: US 10/648,593  
; PRIOR FILING DATE: 2003-08-26  
; NUMBER OF SEQ ID NOS: 571  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 254  
; LENGTH: 822  
; TYPE: PRP  
; ORGANISM: Homo sapiens  
US-11-072-175-254  
Query Match 11.0%; Score 659; DB 7; Length 822;  
Best Local Similarity 27.7%; Pred. No. 1.1e-36;  
Matches 214; Conservative 140; Mismatches 293; Indels 126; Gaps 24;  
QY 381 BEMTLVKPDGTVLHPKDFNH--TDHFSVAI---FTIHRILPDSGVVSVNTVAGMVEK 435  
Db 59 EVRCLKDAAVISWTQGVHGLGNRTVLIGBYLQIKATPRDSGLIYACTASRTVDSETW 118  
QY 436 PFNISKVLPKPLNAPNVIDTGHNFANINISSE--PYFGDPIKSKLLKYPNVNHWBAHQ 493  
Db 119 YFMVNVTDIASSGDEDDDDTDCGAEDFVSENSNNKRAPYNTNTEKMEKRLHAPPAANTVKPR 178  
QY 494 HIQVTNEIVTLNLEPTREYELCVQLVRRGEGEGHPGVRRFTTASIGLPPPGRLNLLP 553

```
Db 179 CPAGNPMTWRLKNGKEK---QEHRIQ-----GKVRNQHSWLIME-----SVVP 223
Qy 554 KSQTLNLTWQIPFSSDDFVFEVERSVQKSDQONIKVPGNLTSV-----LLN 603
Db 224 SDKGNVTCVENEYGSINHVTYLDVVERSPHPIQ-AGLPANASTVVGDDVEFVCKVYS 282
Qy 604 NLHPREQVVRARVNTKAQGESEDLTAWLSDILPQPENIKISNITHSSAVISWTILD 663
Db 283 DAQPHIQWIKHVKNGSKVG-----PDGL-PYLKVLKHSNGINSNAEV----- 324
Qy 664 GYSISITIRYKVGQKNEQHVVDVKIRKNTIYQOLKGLEPETAYQVDFIFAENNINGSSNP 723
Db 325 -LALFNVT-----BADAGEYICK-----VSNYIQGANQ 351
Qy 724 AFSHELVTLPSSQAP-----DLGGGKMLLIALGSGMTCLTVLLAFLILQLKRA----- 775
Db 352 --SAWLTVPKQAPGREKEITASPDYLEIAIYCIYGVFLIACMVVTVILCRMKNTRKPD 409
Qy 776 -----NVQRMAQAFQ-NVREBPVQFNSGT--LALNRKVKNNPD-PTIYPVLDW-- 821
Db 410 PSSQPAVHKLTKRPLRPQVTVSAESSMSMNSNTPLVITRSLSTADTPMLAGVSEYEL 469
Qy 822 -NDIKFQ-----DVIGENFQVLKAR---IKDGLR--MDAAIKRMKEYASKDDH 866
Db 470 PEDPKWEPRDKLTJLKGKLGEGCFQVWMAEAVGIDKDKPEAVTVAVKMLKDDATEKDL 529
Qy 867 RDFAGELEVCLKGLHPHNIINLLGACERHGYLYLAIEYAPHGNLLDFLRKSRVLETDPAF 926
Db 530 SDLVSEMEMKMGHGHKNIINLLGACTQDQPLVIVVEYASKGNREYLRRRPPGMEYSY 589
Qy 927 AIANSTASTLSQQLLHFAADVARGMDVLSOKOFTHRDLAARNILVGENYVAKTADFGLS 986
Db 590 DINRVPEQMTFKDJVSTCYQLARGWEYLSQKCHRDLAARNILVTENNVWKIADFGLA 649
Qy 987 R---GQEVYVKKTMGRLPVRVMAIESLNYSVYVTTNSDVMSYGVLLWEIVSLGGTPYCGMT 1043
Db 650 RDINNIDYKTTNGRLPVKVMAPALFDRVVYTHQSDVMSFGVLWMEVFTLGGSPYGP 709
Qy 1044 CAELYEKLPQGVRLKPLNCDDDEVYDLMRQCRKPYRPSFAQILVSLNRL 1096
Db 710 VEELFKLKEGRHMDKPACTNELYMMRDCWHAVPQSQRPTFKQLVEDLDRIL 762

RESULT 5
US-10-877-346-75
; Sequence 75, Application US/10877346
; Publication No. US20060014153A1
; GENERAL INFORMATION:
; APPLICANT: Gerlach, Valerie L
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Grose, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Leach, Martin D
; APPLICANT: Shinkets, Richard A
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-124
; CURRENT APPLICATION NUMBER: US/10/877,346
; CURRENT FILING DATE: 2004-06-25
; PRIOR APPLICATION NUMBER: US/09/964,956
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/235,631
; PRIOR FILING DATE: 2000-09-27
```

```
; PRIOR APPLICATION NUMBER: 60/235,633
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/235,808
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,064
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,065
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,066
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,135
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/237,434
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/238,321
; PRIOR FILING DATE: 2000-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 75
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Tyrosine
; OTHER INFORMATION: Kinase Consensus Sequence
US-10-877-346-75

Query Match 10.8%; Score 646.5; DB 6; Length 258;
Best Local Similarity 49.4%; Pred. No. 1.6e-36;
Matches 132; Conservative 40; Mismatches 76; Indels 19; Gaps 4;

Qy 830 IGEHFGVGLKARFK-KDGLRMDAAIKRMKEYASKDHRDFAGEVLCKLGHHPHNIINL 888
Db 7 LGEAGFGEVYKGTIKGGGVEVAVKTKEDASEQIEEFLEARLMRKL-DHPNIVKL 65

Qy 889 IGACEHRGYLYLAIEYAPHGNLLDFLRKSRVLETDPAFANSTASTLSQQLHFAADV 948
Db 66 LGVCTEEPLMIVMEYEGDLDLYLKQRPE-----LSLSDLSFALQI 111

Qy 949 ARGMDYLSQKOFTHRDLAARNILVGENYVAKIADFGLSR---GQEVYVKKTMGRLPVRWM 1005
Db 112 ARGMEYLESNFVHRDLAARNCLVGENKTVKIADFGLARDLYDDDYVRKKKSPRLPIRWM 171

Qy 1006 AIESLNYSVYVTTNSDVMSYGVLLWEIVSLGGTPYCGMTCAELYEKLPQGVRLKPLNCDD 1065
Db 172 APESLKDQKFTSKSDVMSFGVLWMEVFTLGGSPYGMSEEVLELYLKGKYLDPQPPNCPD 231

Qy 1066 EYDLMRQCRKPYRPSFAQILVSL 1092
Db 232 EYDLMQLQCAEDPEDRPTFSELVERL 258

RESULT 6
US-11-092-168-9
; Sequence 9, Application US/11092168
; Publication No. US20050277658A1
; GENERAL INFORMATION:
; APPLICANT: Arizona Board of Regents on behalf of The University of Arizona
; APPLICANT: Montigen Pharmaceuticals, Inc.
; APPLICANT: Hurley, Laurence H.
; APPLICANT: Mahadevan, Daruka
; APPLICANT: Han, Haiyong
; APPLICANT: Bearss, David J.
; APPLICANT: Vankayalapati, Hariprasad
; APPLICANT: Bashyam, Sridevi
; APPLICANT: Munoz, Ruben M.
; APPLICANT: Warner, Steven L.
; APPLICANT: Della Croce, Kimiko
; APPLICANT: Von Hoff, Daniel D.
; APPLICANT: Grand, Cory L.
; TITLE OF INVENTION: PROTEIN KINASE INHIBITORS
; FILE REFERENCE: 920214.00003CONT3
```

```
; CURRENT APPLICATION NUMBER: US/11/092,168
; CURRENT FILING DATE: 2005-03-29
; PRIOR APPLICATION NUMBER: US 10/965,313
; US 60/608,529
; US 60/511,486
; US 60/511,489
; PRIOR FILING DATE: 2004-10-14
; 2004-09-09
; 2003-10-14
; 2003-10-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-092-168-9

Query Match      10.5%; Score 633; DB 7; Length 310;
Best Local Similarity 46.9%; Pred. No. 1.6e-35;
Matches 130; Conservative 54; Mismatches 81; Indels 12; Gaps 3;

QY 830 IGEHFGQVLKARIKDGKLRMD-----AATKRMKEYASKDDHRDFACELEVLCKLGH 882
DB 29 IGEHFGQVLA--EAIGLDKDKPRVTYKVAVMKLSKDATEKDLSDLISEMMKMGKH 86

QY 883 PNIINLLGACEHGYLYLAIEYAPHGNLDFLRKSRVLETDPAFAIANSTASTLSSQQL 942
DB 87 KNIINLLGACTDGPYIVIVEYASKGNLREYLQARRPPGLEYSYNPSPHPEQLSSKOLIV 146

QY 943 HPAADVARGMDYLSQKQFHRDLAARNILVGNVYAKIADFGLSRG---QEVYVKTWGR 999
DB 147 SCAYQVARGMEYLASKKCIHRDLAARNVLVTDNVMKIADFGLARDIHHIDYKKTNGR 206

QY 1000 LPVRWMAIESLNSVYTTNSDVMVSGVLLWEIVLSGGTGYCGMTCAELEYEKLPOGVRL 1059
DB 207 LEVWMAPEALPDRIYTHQSDVMVSGVLLWEIVLTGGSPYGPVPEELFKLKEGHRMDK 266

QY 1060 PLNCDDVYDLRMQCRKPEYRPPGFAQILVSLNRLM 1096
DB 267 PSNCTNELYMMRDCHWAVPSQRPTEFKQLVEDLDRIV 303

RESULT 7
US-11-113-424-61
; Sequence 61, Application US/11113424
; Publication No. US20050260713A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli et al.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-225
; CURRENT APPLICATION NUMBER: US/11/113,424
; PRIOR FILING DATE: 2005-04-21
; PRIOR APPLICATION NUMBER: 60/256,704
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/311,590
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/257,314
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/311,613
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/315,617
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/307,506
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: 60/322,358
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/294,075
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/288,153
; PRIOR FILING DATE: 2001-05-02
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 61
; LENGTH: 985
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-11-113-424-61

Query Match      10.2%; Score 612; DB 7; Length 985;
Best Local Similarity 29.2%; Pred. No. 2e-33;
Matches 182; Conservative 104; Mismatches 251; Indels 86; Gaps 23;

QY 544 PPRGLNLLPK-SQTLNLTQPIPPSSE---DDFYVEVER-----SVQK 595
DB 328 PPSAPQNLISNVNETSIVNLEWSP--PQNSGGRPDVSYNLVCKRCSDDLTRCSPGSGVHY 385

QY 586 SDQQN-IKVPGNLTSVLLNHLHPRQYVYVRA-RVN--TKAQGEWSBEDLTAW-TLSDILPP 640
DB 386 SPQNGGLKT---TKVSINDLQAHTNYTFEVAINGVSKQNFQEQDQAVSVTVTTNQAAPS 441

QY 641 QPENIKISNITHSSAVISWTILDGYSISSITIRYKQGNEDQHV-DVKIKNATIIQYQL 699
DB 442 TVTQIQPKETIRHSVSLTWP--EPERANGVILEYEVKYEKQDQNSRYRIVKTAASADI 499

QY 700 KGLEPETAYQVDIFAEN--NIGSSNPAPFSHELVTLPESQAPADLGGKMLLIALLGSAGM 757
DB 500 KGLNELTGTGFVHRARTAGYGFSGPFEFTNTVPSPM----IGEGTPTVLLVSVAGS 555

QY 758 TCLTVLLAFLIILQKRANVORMAQAFQNVREEPVAFQFNSGTLALNRKVKNNPDPTIY- 816
DB 556 IVLVILIAAFVISRRS-----KYSKAKQEADEEK-----HLNQGVKTYVDPFTYE 602

QY 817 -----PVLWDNDIKFQDVIGEGNFGQVLKARIKQDGLR-MDAAIKRMKEYASKDD 865
DB 603 DPNOAVRBFAXEIDASCIEKIEKVGVEGFEVCSGLKVPKREIYVAITKLKAGYTDKQ 662

QY 866 HRDFAGELEVLCKLGHNNIINLLGACEHGYLYLAIEYAPHGNLDFLRKSRVLETDPA 925
DB 663 RRDIFUSEASIMQGP-DHPNIIHLEGVVTCKPVMIIITETMENGSLDAFLRKN-----DGR 716

QY 926 FAIANSTASTLSSQQLHFAADVARGMDYLSQKQFHRDLAARNILVGNVYAKIADFG 985
DB 717 FTVI-----QLVGMRLRGISGMKYLSDMSVHRDLAARNILVNSLNKVSDFGM 766

QY 986 SRGQE-----VYVKTWGRLPVRWMAIESLNSVYTTNSDVMVSGVLLWEIVLSGGTGYC 1040
DB 767 SRVLEDDPEAAVYTRG-GKIPIRWTAPEALAYRKFTSASDVMVSGVIMVMEVMSYGERPYW 825

QY 1041 GMTCAELEYEKLPOGVRLKPLNCDDVYDLRMQCRKPEYRPPGFAQILVSLNRLMEERK 1100
DB 826 DNSNODVIRKAEIEGYRLPPMDPCPIALHQLMDDCWKORSRDKPKFGQIVSMMLDLIRNPN 885

QY 1101 TYVNTTLYEKEFT-YAGIDCSABE 1122
DB 886 SILKRTGLENSRTNTALLDPSSPE 908

RESULT 8
US-11-113-424-59
; Sequence 59, Application US/11113424
; Publication No. US20050260713A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli et al.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-225
; CURRENT APPLICATION NUMBER: US/11/113,424
; PRIOR FILING DATE: 2005-04-21
; PRIOR APPLICATION NUMBER: 60/256,704
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/311,590
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/257,314
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/311,613
; PRIOR FILING DATE: 2001-08-10
```

```

; PRIOR APPLICATION NUMBER: 60/315,617
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/307,506
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: 60/322,358
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/294,075
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/288,153
; PRIOR FILING DATE: 2001-05-02
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 59
; LENGTH: 983
; TYPE: PRF
; ORGANISM: Homo sapiens
US-11-113-424-59

Query Match          10.2%; Score 611.5; DB 7; Length 983;
Best Local Similarity 24.0%; Pred. No. 2.2e-33;
Matches 208; Conservative 111; Mismatches 259; Indels 287; Gaps 32;

QY 284 PYG-CSCATGW--KGLQCNACHPGFY-GPDKLRCS-C--NNGEMCDRFQGLCSPGW- 335
DB 253 PIGKCSNAGYEERFMC-QACRPGFYKALDGNMCKAKCPPHSSQTQEDGSMNCRCENNYF 311
QY 336 -----QGLQCREGIPRMTPK-----IVDLPHIEVNSGK---FNPICKASGW 375
DB 312 RADKPPSMACR---PPSSPRNVISNINETSVIDLWSWPLDTGGRKDVTFNICKCGW 368
QY 376 PLPTNEMTLVKPDGTVLHPKDFNHTDHSVAIFTIHRILPDSGVVCSVNTVAGMVEK 435
DB 369 NIKQCEP-----VFELPR-----QFG-----SVTIL----- 487
QY 436 PFNISVKVLPKPLNAPNVIDTGHNFVAVINISSEPYFGDPIKSKLLYKPNVHYEAWQHI 495
DB 380 -----VFELPR-----QFG-----SVTIL----- 487
QY 496 QVTNEIVTLNLYPRTEYELCVLVRGEGGHPGVRRTTASI--GLPPRGLNLLP 553
DB 399 -LTNTVTVTDLAHTNTYTFEIDAV---NGVSELSSPPRQFAAVSITTNQAAPSPVLTIK 444
QY 554 KSQTTLN---LTWQPIFPSSDDFVVEVERRSVQSDQONIKVPGNLTSLVLLNLHPREQ 610
DB 445 KDRTSRNSISLSWQ--PEHPNGIILDYEVKYEQEQT----- 482
QY 611 YVVRARVNTKAQGEWSEDLTAWTSLDILPPQENIKISNITHSSAVISWTILDGYSISSI 670
DB 483 -----VFELPR-----QFG-----SVTIL----- 487
QY 671 TIRYKVOGKNEQHVVDVKIKNATIIQYQLKGLPETAYQVDIFAEENNIGSSNPAFSHEL 730
DB 488 -----RARGTN-----VTISLKPDTIYVFOIRARTAAAGYGTNSRKFEFE 527
QY 731 TLPESQAPADLGG--GKMLLIALGSAGTCTLVLLAFLLIQLKRVANVQRMQAQFQNV 788
DB 528 TSPDS---FSISGESQVVMIAISAVALILLTVVIYVLI-----GRFCGYSKRG 575
QY 789 REPAVQFNSGTALNRKYNPNPDPTIY-----PVLDMNDIKPDQVIGEGFQ 837
DB 576 ADEKRLHFGNHLKL-PGLRTVVDPTHTYEDPTQAVHEFAKELDATNISIDKVVGAGEFGE 634
QY 838 VLKARIKQDGLR-MDAAIKRMKEYASKDHRDPAGELEVLCGLHHPNIIINLLGACEHRG 896
DB 635 VCSGRKLKPSKKEISVAIKLVGYTEKQRDRFLGEASIMQGF-DHPNIIIRLEGVYVTKSK 693
QY 897 YLYLAIEYAPHGNLLDFLRKSRVLETDPAFAFANGSTASLSQQLHFAADVARGMDVLS 956
DB 694 PWMIVTEWENGLSDFLRKH-----DAQFTVI-----QLVGLMRGIASGMKYL 738
QY 957 QKQFTHRDLAARNILVGENYAKIADFLGRQGE-----VYKTKMGRLPVWRMAIESLN 1011
DB 739 DMGYVHRDLAARNILNSLVCKVSDFLGRVLEDDPEAAVYTRG-GKIPRWTSPEAIA 797

```

```

QY 1012 YSVYTTNSDVSYSYVLLMEIVSLGCTPYCGMTCAELYEKLPQGYRLEKPLNCDEVDLM 1071
DB 798 YRKFTSASDVSYSYVLLMEIVSLGCTPYCGMTCAELYEKLPQGYRLEKPLNCDEVDLM 857
QY 1072 RQWREKPYRPSFAQILVSLNRLM 1096
DB 858 LDCWQKDRNNRPKFEQIVSILDKLI 882

```

## RESULT 9

```

US-11-203-251A-78
; Sequence 78, Application US/11203251A
; Publication No. US20060039904A1
; GENERAL INFORMATION:
; APPLICANT: MedImmune Inc.
; TITLE OF INVENTION: EPH RECEPTOR FC VARIANTS WITH ENHANCED ANTIBODY DEPENDENT
; CELL-MEDIATED CYTOTOXICITY ACTIVITY
; FILE REFERENCE: AE702US
; CURRENT APPLICATION NUMBER: US/11/203,251A
; CURRENT FILING DATE: 2005-08-15
; PRIOR APPLICATION NUMBER: 60/601,634
; PRIOR FILING DATE: 2004-08-16
; PRIOR APPLICATION NUMBER: 60/608,852
; PRIOR FILING DATE: 2004-09-13
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: Patent In version 3.3
; SEQ ID NO 78
; LENGTH: 983
; TYPE: PRF
; ORGANISM: Homo sapiens
US-11-203-251A-78

```

```

Query Match          10.2%; Score 611.5; DB 7; Length 983;
Best Local Similarity 24.0%; Pred. No. 2.2e-33;
Matches 208; Conservative 111; Mismatches 259; Indels 287; Gaps 32;

QY 284 PYG-CSCATGW--KGLQCNACHPGFY-GPDKLRCS-C--NNGEMCDRFQGLCSPGW- 335
DB 253 PIGKCSNAGYEERFMC-QACRPGFYKALDGNMCKAKCPPHSSQTQEDGSMNCRCENNYF 311
QY 336 -----QGLQCREGIPRMTPK-----IVDLPHIEVNSGK---FNPICKASGW 375
DB 312 RADKPPSMACR---PPSSPRNVISNINETSVIDLWSWPLDTGGRKDVTFNICKCGW 368
QY 376 PLPTNEMTLVKPDGTVLHPKDFNHTDHSVAIFTIHRILPDSGVVCSVNTVAGMVEK 435
DB 369 NIKQCEP-----VFELPR-----QFG-----SVTIL----- 487
QY 436 PFNISVKVLPKPLNAPNVIDTGHNFVAVINISSEPYFGDPIKSKLLYKPNVHYEAWQHI 495
DB 380 -----VFELPR-----QFG-----SVTIL----- 487
QY 496 QVTNEIVTLNLYPRTEYELCVLVRGEGGHPGVRRTTASI--GLPPRGLNLLP 553
DB 399 -LTNTVTVTDLAHTNTYTFEIDAV---NGVSELSSPPRQFAAVSITTNQAAPSPVLTIK 444
QY 554 KSQTTLN---LTWQPIFPSSDDFVVEVERRSVQSDQONIKVPGNLTSLVLLNLHPREQ 610
DB 445 KDRTSRNSISLSWQ--PEHPNGIILDYEVKYEQEQT----- 482
QY 611 YVVRARVNTKAQGEWSEDLTAWTSLDILPPQENIKISNITHSSAVISWTILDGYSISSI 670
DB 483 -----VFELPR-----QFG-----SVTIL----- 487
QY 671 TIRYKVOGKNEQHVVDVKIKNATIIQYQLKGLPETAYQVDIFAEENNIGSSNPAFSHEL 730
DB 488 -----RARGTN-----VTISLKPDTIYVFOIRARTAAAGYGTNSRKFEFE 527
QY 731 TLPESQAPADLGG--GKMLLIALGSAGTCTLVLLAFLLIQLKRVANVQRMQAQFQNV 788
DB 528 TSPDS---FSISGESQVVMIAISAVALILLTVVIYVLI-----GRFCGYSKRG 575

```



US-11-148-770-31

```
Query Match      10.1%; Score 608; DB 7; Length 976;
Best Local Similarity 23.4%; Pred. No. 3.7e-33;
Matches 256; Conservative 121; Mismatches 326; Indels 390; Gaps 40;

QY 206 LIVRCEAQKQWPECNHLCTACMNGVCHDTEGECICPPGFMGRTCCEKACELH----- 258
DB |||||-----VGDEIRLLCT-----DPGFVKWTFEILDENKQNEW 82

QY 259 -----TFGRTCKERCSSGQGGCKS--YVFCLPDPYGCSCATGWKGLQCNEACHPGFYGP-- 309
DB |||||-----TFGRTCKERCSSGQGGCKS--YVFCLPDPYGCSCATGWKGLQCNEACHPGFYGP-- 309

QY 83 ITEKAEATNTGKYTCNKHLSNSIYVF-VRDPAKLFLV-----DRSLYGRKD 129
DB |||||-----ITEKAEATNTGKYTCNKHLSNSIYVF-VRDPAKLFLV-----DRSLYGRKD 129

QY 310 -DCKLRCSNNGEMCD-RFOGC-----LCSPFGWQG 337
DB |||||-----DCKLRCSNNGEMCD-RFOGC-----LCSPFGWQG 337

QY 130 NDTLVRCPLTDPVNTNYSLSKGCQKPLPKDLRFIPDPKAGIMIKSVGRAYHRLC----- 183
DB |||||-----NDTLVRCPLTDPVNTNYSLSKGCQKPLPKDLRFIPDPKAGIMIKSVGRAYHRLC----- 183

QY 338 LOC--EREGIPRMTPKIVDLPDHIENVSGKFNPICKA-----SGWPLPTNEEMTLV- 386
DB |||||-----EREGIPRMTPKIVDLPDHIENVSGKFNPICKA-----SGWPLPTNEEMTLV- 386

QY 184 LHCSVDQEGSKVLSKFI-----LKVRPAKAVPVVSVSKASYLLRREGGEFTVTC 233
DB |||||-----LHCSVDQEGSKVLSKFI-----LKVRPAKAVPVVSVSKASYLLRREGGEFTVTC 233

QY 387 -----KPDGTVL-----HPKDFNHTDHFSAIFTIHRILPPDSGVWVC 424
DB |||||-----KPDGTVL-----HPKDFNHTDHFSAIFTIHRILPPDSGVWVC 424

QY 234 TIKDVSSSVYSTWKRENSQTKLQEKYNSWHHGDFNYERQ---ATLTISSARVNDSGVFM 290
DB |||||-----TIKDVSSSVYSTWKRENSQTKLQEKYNSWHHGDFNYERQ---ATLTISSARVNDSGVFM 290

QY 425 SVNTVAG-----MVEKPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPFYFGDP 475
DB |||||-----MVEKPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPFYFGDP 475

QY 291 YANNTFGSANVTTLTLEVDKGF---INIFPM-INTTVFVNDGENVDLI----- 334
DB |||||-----YANNTFGSANVTTLTLEVDKGF---INIFPM-INTTVFVNDGENVDLI----- 334

QY 476 IKSCKLLYPVNHBEAWQHIQVTNEIVTLNYPELTEYELCVQLVRRGEGGHPGVR 535
DB |||||-----VEYEF-----PKPEHQOYIMNR----- 353

QY 536 FTTASIGLPPRGLNLLPKSQTLNLTWQIPFSSSEDDFYVEVERRSVQSKDQONIK-VP 594
DB |||||-----FTTASIGLPPRGLNLLPKSQTLNLTWQIPFSSSEDDFYVEVERRSVQSKDQONIK-VP 594

QY 354 -----TFTDKWED-YPKSENEISNI-----RYVSELHLTKLKGTE 386
DB |||||-----TFTDKWED-YPKSENEISNI-----RYVSELHLTKLKGTE 386

QY 595 GNLTSVLLNHLHPRQYVVRARVNTKAQGEWSEDLTAWTSLDILPPQENIKISNIHSS 654
DB |||||-----GNLTSVLLNHLHPRQYVVRARVNTKAQGEWSEDLTAWTSLDILPPQENIKISNIHSS 654

QY 387 GGTTFYLSNDSVNAIAFNVTNTKPEILTYDRLVNGMLQCVAAAGPPE----- 435
DB |||||-----GGTTFYLSNDSVNAIAFNVTNTKPEILTYDRLVNGMLQCVAAAGPPE----- 435

QY 655 AVISWTILDG---YSISITIRYKVOGKNEQDQVVKIKNAT-----IIQYOLKGLE 703
DB |||||-----AVISWTILDG---YSISITIRYKVOGKNEQDQVVKIKNAT-----IIQYOLKGLE 703

QY 436 PTIDWYFCGTQORCSASVLP-----VDVQTLNLSGGPPFGKLVVQSSIDSSA 482
DB |||||-----PTIDWYFCGTQORCSASVLP-----VDVQTLNLSGGPPFGKLVVQSSIDSSA 482

QY 704 PETAYQVDIFAENNISSNPAFS-----HELVTLPSQAPADLGGGRMLLIA 750
DB |||||-----PETAYQVDIFAENNISSNPAFS-----HELVTLPSQAPADLGGGRMLLIA 750

QY 483 FXHNGTVECKAYNDVGKTSAYENFAFGKNNKEQIHPHTLFT-----PLLIIG 528
DB |||||-----FXHNGTVECKAYNDVGKTSAYENFAFGKNNKEQIHPHTLFT-----PLLIIG 528

QY 751 ILGSGAGMTCLTVLLAFLLIQLKRVNQRMAQAQFQNVREBPAQFNSGTLALNRKVQN 810
DB |||||-----ILGSGAGMTCLTVLLAFLLIQLKRVNQRMAQAQFQNVREBPAQFNSGTLALNRKVQN 810

QY 529 FVIVAGMCIIVMLITYKLOKPMYEVQMKVVEEING-----NN 567
DB |||||-----FVIVAGMCIIVMLITYKLOKPMYEVQMKVVEEING-----NN 567

QY 811 ---PDPITVFPV-LDW---NDIKFQDVIDEGNFGVOLKAR---IKDGLRMDAAIKRMK 858
DB |||||-----PDPITVFPV-LDW---NDIKFQDVIDEGNFGVOLKAR---IKDGLRMDAAIKRMK 858

QY 568 YVIIDPTQLPYDHHKEFPNRLUSFGKTLGAGAFKVVEATATGLKSDA-AMTAVAKMLK 626
DB |||||-----YVIIDPTQLPYDHHKEFPNRLUSFGKTLGAGAFKVVEATATGLKSDA-AMTAVAKMLK 626

QY 859 EYAKSDDRDFAGELEVLCKLGHFHNINLLGACEHRGYLYLAIYAPHGNLLDLFRKSR 918
DB |||||-----EYAKSDDRDFAGELEVLCKLGHFHNINLLGACEHRGYLYLAIYAPHGNLLDLFRKSR 918

QY 627 PSAHLTEREALMSCLKVLSYLGHNHNIIVNLLGACTIGGPTLVTIYCCVCGDLLNLFRRKR 686
DB |||||-----PSAHLTEREALMSCLKVLSYLGHNHNIIVNLLGACTIGGPTLVTIYCCVCGDLLNLFRRKR 686

QY 919 ----- 918
DB ----- 918

QY 687 DSFICSKQBDHAEALYKNLHLSKSSCSDSTNEYMDMKPGVSYVVPYTKADKRRSVRIGS 746
DB -----DSFICSKQBDHAEALYKNLHLSKSSCSDSTNEYMDMKPGVSYVVPYTKADKRRSVRIGS 746

QY 919 VLETPAPAIANSTASTSSQOLLHFAADVARGMDVLSOKQFIHRDLAARNILVGENYVA 978
DB |||||-----VLETPAPAIANSTASTSSQOLLHFAADVARGMDVLSOKQFIHRDLAARNILVGENYVA 978

QY 747 YIERDVTPTAIMEDELALDLELLSFSYQVAKGMFLASKNCIHRDLAARNILLTHGRIT 806
DB |||||-----YIERDVTPTAIMEDELALDLELLSFSYQVAKGMFLASKNCIHRDLAARNILLTHGRIT 806

QY 979 KIADFGLSR---GQEVYVKKTWGRLPVRMAJTESLNSYVYTTNSDVWSYGLLWELFISLG 1035
DB |||||-----KIADFGLSR---GQEVYVKKTWGRLPVRMAJTESLNSYVYTTNSDVWSYGLLWELFISLG 1035

QY 807 KICDFGLARDIKNDSNYVVKGNARLPVKWMAPEISFNCVYTPESDVWSYGIFLWELFISLG 866
DB |||||-----KICDFGLARDIKNDSNYVVKGNARLPVKWMAPEISFNCVYTPESDVWSYGIFLWELFISLG 866
```

```
QY 1036 GTPYCGMTC-AELYEKLPOGYRLEKPLNCDDEVYDLMRQCHREKPYERPSFAQILVSLNR 1094
DB |||||-----GTPYCGMTC-AELYEKLPOGYRLEKPLNCDDEVYDLMRQCHREKPYERPSFAQILVSLNR 1094

QY 867 SSPYGPMPVDKFKYKMIKEGFRMLSPHEAPAEYMTIMKTCDADPLKRPYFKQIVQLIEK 926
DB |||||-----SSPYGPMPVDKFKYKMIKEGFRMLSPHEAPAEYMTIMKTCDADPLKRPYFKQIVQLIEK 926

QY 1095 MLEERKTYVNTTL 1107
DB |||||-----MLEERKTYVNTTL 1107

QY 927 QISESTNHIYSL 939
DB |||||-----QISESTNHIYSL 939

RESULT 12
US-11-177-894-15
; Sequence 15, Application US/11177894
; Publication No. US20060040292A1
; GENERAL INFORMATION:
; APPLICANT: West, et al.
; TITLE OF INVENTION: Tumor Markers and Uses Thereof
; FILE REFERENCE: 2002850-0048
; CURRENT APPLICATION NUMBER: US/11/177,894
; CURRENT FILING DATE: 2005-07-08
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 15
; LENGTH: 976
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: KIT amino acid sequence
US-11-177-894-15
```

```
Query Match      10.1%; Score 608; DB 7; Length 976;
Best Local Similarity 23.4%; Pred. No. 3.7e-33;
Matches 256; Conservative 121; Mismatches 326; Indels 390; Gaps 40;

QY 206 LIVRCEAQKQWPECNHLCTACMNGVCHDTEGECICPPGFMGRTCCEKACELH----- 258
DB |||||-----VGDEIRLLCT-----DPGFVKWTFEILDENKQNEW 82

QY 259 -----TFGRTCKERCSSGQGGCKS--YVFCLPDPYGCSCATGWKGLQCNEACHPGFYGP-- 309
DB |||||-----TFGRTCKERCSSGQGGCKS--YVFCLPDPYGCSCATGWKGLQCNEACHPGFYGP-- 309

QY 83 ITEKAEATNTGKYTCNKHLSNSIYVF-VRDPAKLFLV-----DRSLYGRKD 129
DB |||||-----ITEKAEATNTGKYTCNKHLSNSIYVF-VRDPAKLFLV-----DRSLYGRKD 129

QY 310 -DCKLRCSNNGEMCD-RFOGC-----LCSPFGWQG 337
DB |||||-----DCKLRCSNNGEMCD-RFOGC-----LCSPFGWQG 337

QY 130 NDTLVRCPLTDPVNTNYSLSKGCQKPLPKDLRFIPDPKAGIMIKSVGRAYHRLC----- 183
DB |||||-----NDTLVRCPLTDPVNTNYSLSKGCQKPLPKDLRFIPDPKAGIMIKSVGRAYHRLC----- 183

QY 338 LOC--EREGIPRMTPKIVDLPDHIENVSGKFNPICKA-----SGWPLPTNEEMTLV- 386
DB |||||-----EREGIPRMTPKIVDLPDHIENVSGKFNPICKA-----SGWPLPTNEEMTLV- 386

QY 184 LHCSVDQEGSKVLSKFI-----LKVRPAKAVPVVSVSKASYLLRREGGEFTVTC 233
DB |||||-----LHCSVDQEGSKVLSKFI-----LKVRPAKAVPVVSVSKASYLLRREGGEFTVTC 233

QY 387 -----KPDGTVL-----HPKDFNHTDHFSAIFTIHRILPPDSGVWVC 424
DB |||||-----KPDGTVL-----HPKDFNHTDHFSAIFTIHRILPPDSGVWVC 424

QY 234 TIKDVSSSVYSTWKRENSQTKLQEKYNSWHHGDFNYERQ---ATLTISSARVNDSGVFM 290
DB |||||-----TIKDVSSSVYSTWKRENSQTKLQEKYNSWHHGDFNYERQ---ATLTISSARVNDSGVFM 290

QY 425 SVNTVAG-----MVEKPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPFYFGDP 475
DB |||||-----MVEKPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPFYFGDP 475

QY 291 YANNTFGSANVTTLTLEVDKGF---INIFPM-INTTVFVNDGENVDLI----- 334
DB |||||-----YANNTFGSANVTTLTLEVDKGF---INIFPM-INTTVFVNDGENVDLI----- 334

QY 476 IKSCKLLYPVNHBEAWQHIQVTNEIVTLNYPELTEYELCVQLVRRGEGGHPGVR 535
DB |||||-----VEYEF-----PKPEHQOYIMNR----- 353

QY 536 FTTASIGLPPRGLNLLPKSQTLNLTWQIPFSSSEDDFYVEVERRSVQSKDQONIK-VP 594
DB |||||-----FTTASIGLPPRGLNLLPKSQTLNLTWQIPFSSSEDDFYVEVERRSVQSKDQONIK-VP 594

QY 354 -----TFTDKWED-YPKSENEISNI-----RYVSELHLTKLKGTE 386
DB |||||-----TFTDKWED-YPKSENEISNI-----RYVSELHLTKLKGTE 386

QY 595 GNLTSVLLNHLHPRQYVVRARVNTKAQGEWSEDLTAWTSLDILPPQENIKISNIHSS 654
DB |||||-----GNLTSVLLNHLHPRQYVVRARVNTKAQGEWSEDLTAWTSLDILPPQENIKISNIHSS 654

QY 387 GGTTFYLSNDSVNAIAFNVTNTKPEILTYDRLVNGMLQCVAAAGPPE----- 435
DB |||||-----GGTTFYLSNDSVNAIAFNVTNTKPEILTYDRLVNGMLQCVAAAGPPE----- 435

QY 655 AVISWTILDG---YSISITIRYKVOGKNEQDQVVKIKNAT-----IIQYOLKGLE 703
DB |||||-----AVISWTILDG---YSISITIRYKVOGKNEQDQVVKIKNAT-----IIQYOLKGLE 703

QY 436 PTIDWYFCGTQORCSASVLP-----VDVQTLNLSGGPPFGKLVVQSSIDSSA 482
DB |||||-----PTIDWYFCGTQORCSASVLP-----VDVQTLNLSGGPPFGKLVVQSSIDSSA 482
```

QY 704 PETAYQVDIFABNNIGSSNPAPS-----HELVTLPEAQADLGGKMLLIA 750  
DB : : : : :  
483 FXHNGTVECKAYNDVGKTSAYFNFAFGKGNKEQIHPHTLFT-----PLLIIG 528  
QY 751 ILGSGMTCLTVLLAFLIILQKRANVQRMAQAQFQNVREBPAQFNSGTLANRKRKN 810  
DB : : : : :  
529 FVIVAGMCIIVMILTYKYLQKPMYEVOMKVVEING-----NN 567  
QY 811 ---PDPTIYPV-LDW-----NDIKFQDVIPEGNGOVLKAR-----IKOGLRMDAAIKRMK 858  
DB : : : : :  
568 YVYIDPTQLPYDHKEFPNRLSFGKTLGAGAFKVEATAYGLIKSDA-AMTVAVKMLK 626  
QY 859 EVASKDHRDFAGELEVLCKLGHHPNIINLIGACBHRGYLAIYAPHGNLLDFLRKR 918  
DB : : : : :  
627 PSAHLTEREALMSKLVSLGNHNIIVNLLGACTIGGPTLVITEYCCYGDLLNLRKR 686  
QY 919 ----- 918  
DB : : : : :  
687 DSFICKQBDHAAALYKNLLHSSKSSSDSTNEYMDMKPGVSYVYPTKADRRSVRIGS 746  
QY 919 VLETDPAFAIANSTASTLSSQOLLHFAADVARGMDYLSQKQFIHRDLAARNILVGENYVA 978  
DB : : : : :  
747 YIERDVTTPAIMEDELALDLELLSFSYQVAKGMAFLASKNCIHRDLAARNILLTHGRIT 806  
QY 979 KIADFGLSR---GQEVYVKTMGRLPVRMMAIESLNYSVYTTNSDVMSYGVLLWEIVSLG 1035  
DB : : : : :  
807 KICDFGLARDIKNDSNVYVKGNAFLPVKMAPESEIFNCVYTFESDVMSYGVLLWEIVSLG 866  
QY 1036 GTPYCGMTC-AELYEKLPGYRLEKPLNCDDEVYDLMROCWEKPERPSPAQILVSLNR 1094  
DB : : : : :  
867 SSYPGMPVDSKPYKMIKEGFRMLSPHAPAEYIMDKTCDADPLKRPFTFKQIVOLIEK 926  
QY 1095 MLEERKTYVNTTL 1107  
DB : : : : :  
927 QISESTNHIYSNL 939

RESULT 13  
US-11-177-894-16  
; Sequence 16, Application US/11177894  
; Publication No. US20060040292A1  
; GENERAL INFORMATION:  
; APPLICANT: West, et al.  
; TITLE OF INVENTION: Tumor Markers and Uses Thereof  
; FILE REFERENCE: 2002850-0048  
; CURRENT APPLICATION NUMBER: US/11/177,894  
; CURRENT FILING DATE: 2005-07-08  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 16  
; LENGTH: 976  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: KIT amino acid sequence  
US-11-177-894-16

Query Match 10.1%; Score 608; DB 7; Length 976;  
Best Local Similarity 23.4%; Pred. No. 3.7e-33;  
Matches 256; Conservative 121; Mismatches 326; Indels 390; Gaps 40;

QY 206 LIVRCEAKWGPCNHLCTACMNGVCHEDTGECICPGFMGRTCERACELH-----258  
DB : : : : :  
46 LIVR-----VGDEIRLLCT-----DPGFVKWTFEILDETNENKQNEW 82  
QY 259 -----TFGRTKERCSCGQEGCKS---YVFCPLDPYCGCATGWKGLQCNEACHPGYPGP--309  
DB : : : : :  
83 ITEKABATWTKYCTCNKHLGNSIYVF-VROPKFLV-----DRSLYGRKD 129  
QY 310 -PCKLRCSNCGEMCD-RPQGC-----LCSPGHQG 337  
DB : : : : :  
130 NDTLVRCPLTDPVNTNLSLKGCQKGPLPKDLRFIPDPKAGIMIKSVKRAYHRLC-----183

QY 338 LQC--EREGIPRMTPKIVDLPDHIENVSGKFNPICKA-----SGMPLPTNEEMTLV- 386  
DB : : : : :  
184 LHCSVDQSGKSVLSEKFI-----LKVRPAFAKAVPVSVSKASYLLREGGEFTVTC 233  
QY 387 -----KPDGTVL-----HPKDFNHTDHPFSAVFIPIHRLIPDGSVWVC 424  
DB : : : : :  
234 TIKDYSSSVYSTWKRENSQTKLQEKYNWHHGDFNYERQ---ATLTISARVNDSGVFMC 290  
QY 425 SVNTVAG-----MYEKPENISVKVLPKPLNAPNVIDTGHNFANINISSEPEYFGDGP 475  
DB : : : : :  
291 YANFTFGSANVTTLLEVVDKGF---INIFPM-INTVFVNDGENVDLI-----334  
QY 476 IKSXKLLYKPVNHYBAWQHIOVTNEIVTLNLYLEPRTEYELCVQLVRRGEGGEGHPVPR 535  
DB : : : : :  
335 -----VEYEA-----PKPEHQOQIYMNK-----353  
QY 536 FTASIGLPPRGLNLLPKSQTTNLNTWQIPFSSSEDDFYVEVERRSVQSKSQOQNIK-VP 594  
DB : : : : :  
354 -----TFTDKWED-YPKSENESNI-----RYVSELHLTRLKGTE 386  
QY 595 GNLTSVLLNNLHPRQYVVRARVNTKAQGEWSEDLTAWTSLDILPPQPENIKISNITHSS 654  
DB : : : : :  
387 GGTYTFVNSDVNNAIAIPNVYVNTKPELITYDRLVNGMLQCVAAAGFPE-----435  
QY 655 AVISWTILDG---YSISITIRYKVOGKNEDQHVVDKIKNAT-----IIYQOLKGLE 703  
DB : : : : :  
436 PTIDWYFCPGTEQRCASVLP-----VDVQTLNSSSGPPFGKLVVQSSIDSSA 482  
QY 704 PETAYQVDIFABNNIGSSNPAPS-----HELVTLPEAQADLGGKMLLIA 750  
DB : : : : :  
483 FKHNCTVECKAYNDVGKTSAYFNFAFGKGNKEQIHPHTLFT-----PLLIIG 528  
QY 751 ILGSGMTCLTVLLAFLIILQKRANVQRMAQAQFQNVREBPAQFNSGTLANRKRKN 810  
DB : : : : :  
529 FVIVAGMCIIVMILTYKYLQKPMYEVOMKVVEING-----NN 567  
QY 811 ---PDPTIYPV-LDW-----NDIKFQDVIPEGNGOVLKAR-----IKOGLRMDAAIKRMK 858  
DB : : : : :  
568 YVYIDPTQLPYDHKEFPNRLSFGKTLGAGAFKVEATAYGLIKSDA-AMTVAVKMLK 626  
QY 859 EVASKDHRDFAGELEVLCKLGHHPNIINLIGACBHRGYLAIYAPHGNLLDFLRKR 918  
DB : : : : :  
627 PSAHLTEREALMSKLVSLGNHNIIVNLLGACTIGGPTLVITEYCCYGDLLNLRKR 686  
QY 919 ----- 918  
DB : : : : :  
687 DSFICKQBDHAAALYKNLLHSSKSSSDSTNEYMDMKPGVSYVYPTKADRRSVRIGS 746  
QY 919 VLETDPAFAIANSTASTLSSQOLLHFAADVARGMDYLSQKQFIHRDLAARNILVGENYVA 978  
DB : : : : :  
747 YIERDVTTPAIMEDELALDLELLSFSYQVAKGMAFLASKNCIHRDLAARNILLTHGRIT 806  
QY 979 KIADFGLSR---GQEVYVKTMGRLPVRMMAIESLNYSVYTTNSDVMSYGVLLWEIVSLG 1035  
DB : : : : :  
807 KICDFGLARDIKNDSNVYVKGNAFLPVKMAPESEIFNCVYTFESDVMSYGVLLWEIVSLG 866  
QY 1036 GTPYCGMTC-AELYEKLPGYRLEKPLNCDDEVYDLMROCWEKPERPSPAQILVSLNR 1094  
DB : : : : :  
867 SSYPGMPVDSKPYKMIKEGFRMLSPHAPAEYIMDKTCDADPLKRPFTFKQIVOLIEK 926  
QY 1095 MLEERKTYVNTTL 1107  
DB : : : : :  
927 QISESTNHIYSNL 939

RESULT 14  
US-11-177-894-18  
; Sequence 18, Application US/11177894  
; Publication No. US20060040292A1  
; GENERAL INFORMATION:  
; APPLICANT: West, et al.  
; TITLE OF INVENTION: Tumor Markers and Uses Thereof

```

; FILE REFERENCE: 2002850-0048
; CURRENT APPLICATION NUMBER: US/11/177,894
; CURRENT FILING DATE: 2005-07-08
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 976
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: KIT amino acid sequence
US-11-177-894-18

Query Match          10.1%; Score 608; DB 7; Length 976;
Best Local Similarity 23.4%; Pred. No. 3.7e-33;
Matches 256; Conservative 121; Mismatches 326; Indels 390; Gaps 40;

Qy 206 L I V R R C E A Q K W G P E C N H L C T A C M N N G V C H E D T G E C I C P P G P M G T C K A C E L H ----- 258
Db 46 L I V R ----- V G D E I R L L C T ----- D P G F V K W T F E I L D E T N K N Q E W 82

Qy 259 ----- T F G T C K E R C S G Q E G C K S -- Y V F C L P D P Y G C S C A T G K W G L Q C N E A C H P G F Y G P -- 309
Db 83 I T E K A E A T N G K Y T C T N K H G L S N I Y V F - V R D P A K L F I V ----- D R S L Y K G E D 129

Qy 310 - D C K L R C S C N N G E M C D - R F O G C ----- L C S P G W Q G 337
Db 130 N D T L V R C P L D P E V T N Y S L A G C Q G K P L P K O L R F I P D P K A G I M K S V K R A V H R L C ----- 183

Qy 338 L Q C -- E R E G I P R M T P K I V D L P D H L E V N S G K E N P I C K A ----- S G W P L P T N E E M T L V - 386
Db 184 L H C S V D Q E G K S V L S E K P I ----- L K V R P A K A V P V V S K A S Y L L R E G E E F T V T C 233

Qy 387 ----- K P D G T V L ----- H P K F N H T D H F S V A I F T I H R I L P D S G V M V C 424
Db 234 T I K D V S S V Y T W K R E N S Q T K L Q E K Y N S W H G D F N Y E R Q --- A T L T I S S A R V N D S G V F M C 290

Qy 425 S V N T V A G ----- M V E K P F N I S V K V L P K P L N A P N V I D T G H N F A V I N I S S E P Y F G D G P 475
Db 291 Y A N N T F G S A N V T T L E V D K G F --- I N I F M - I N T T V F V N D G E N V D L I ----- 334

Qy 476 I K S K K L L Y K P N V H Y E A M O H I Q V T N E I V T L N Y L E P R T E Y E L C V Q L V R R E G G E G H G P G P V R R 535
Db 335 ----- V E Y E A R ----- P R P E H Q O W I Y M N R ----- 353

Qy 536 F T T A S I G L P P P R G L N L L P K S Q T T L N L T W O P I F P S S E D D F Y V E V E R R S V O K S D Q O N I K - V P 594
Db 354 ----- T F T D K W E D - Y P K S E N S N I ----- R V V S E L H L T R L A G T E 386

Qy 595 G N L T S V L L N N L H P R E Q V V R A R V N T K A G E S E D L T A W T L S D I L P P O P E N I K I S N I T H S S 654
Db 387 G G T Y T F L V S N S D V N A A I A F N Y V N T K P E I L T Y D R L V N G M L Q C V A A G P P E ----- 435

Qy 655 A V I S W T I L D G --- Y S I S S I T I R Y K V Q G K N E D Q H V D V K I K N A T ----- I I O Y Q L A G L E 703
Db 436 P T I D W Y F C P G T E Q R C S A S V L P ----- V D V Q T L N S S G P P K G L V A Q S I D S S A 482

Qy 704 P E T A Y Q V D I F A E N N I G S N P A F S ----- H E L V T L P S Q A P A D I G G K M L L I A 750
Db 483 F K H G T V E K A Y N D V G K T S A Y F N F A P K G N N K E Q I H P H T L P T ----- P L L I G 528

Qy 751 I L G S A G M T C L T V L L A F L I I L Q L K R A N V O R M A Q A F Q N V R E P A V Q F N S G T T A L N R K V K N N 810
Db 529 F V I V A G M M C I V M I L T Y K Y L Q K P M Y E Q V K V V E E I N G ----- N N 567

Qy 811 --- P D P T I Y P V - L D W --- N D I K F O D V I G E N F G O V L K A R --- I K X D G L R M D A A I K R M K 858
Db 568 Y V I D P T Q L P Y D H K W E F P R N L S F G K T I L G A F G K V V E A T A Y G L I K S D A - A M T V A V M K L 626

Qy 859 E Y A S K D P H R D F A G E L V C K L G H H P N I I N L G A C E H R G Y L V L A I E Y A P H G N L D P L R K S R 918
Db 627 P S A H L T R E A L M S E L K V L S Y L G N H M T V N L I G A T T I G P T L V I T E Y C Y G P L A N F L R K R 686

```





**This Page Blank (uspto)**